Journal of Computational Mathematics, Vol.23, No.3, 2005, 305–320.

ON THE MINIMAL NONNEGATIVE SOLUTION OF NONSYMMETRIC ALGEBRAIC RICCATI EQUATION *1)

Xiao-xia Guo Zhong-zhi Bai
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

We study perturbation bound and structured condition number about the minimal nonnegative solution of nonsymmetric algebraic Riccati equation, obtaining a sharp perturbation bound and an accurate condition number. By using the matrix sign function method we present a new method for finding the minimal nonnegative solution of this algebraic Riccati equation. Based on this new method, we show how to compute the desired M-matrix solution of the quadratic matrix equation $X^2 - EX - F = 0$ by connecting it with the nonsymmetric algebraic Riccati equation, where E is a diagonal matrix and F is an M-matrix.

Mathematics subject classification: 65F10, 65F15, 65N30.
Key words: Nonsymmetric algebraic Riccati equation, Minimal nonnegative solution, Matrix sign function, Quadratic matrix equation.

1. Introduction

In this paper, we will mainly study the nonsymmetric *algebraic Riccati equation* (ARE)

$$XCX - XD - AX + B = 0, \tag{1}$$

where A, B, C, D are given real matrices of sizes $m \times m$, $m \times n$, $n \times m$ and $n \times n$, respectively. To this end, let us define two $(m+n) \times (m+n)$ matrices H and K as follows:

$$H = \begin{pmatrix} D & C \\ -B & -A \end{pmatrix}, \qquad K = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix}. \tag{2}$$

We will focus on the exploration of the minimal nonnegative solution of the ARE(1) by making use of the invariant subspace of the matrix H when K is a nonsingular M-matrix.

We have noticed that sensitivity analysis about other types of algebraic Riccati equations were studied in depth in [17, 18, 10, 6], and direct methods about the linear matrix equations, the special cases of the algebraic Riccati equations, were presented in detail in [8, 9].

This paper is organized as follows. After reviewing some basic notations and results associated with the nonsymmetric ARE(1) in section 2, we give a perturbation bound for the minimal nonnegative solution of the ARE(1) in section 3. A structured condition number is derived mathematically and verified numerically in section 4. Then, we present a matrix sign function method for finding the minimal nonnegative solution in section 5; this method can also be used to find the desired M-matrix solution of the quadratic matrix equation $X^2 - EX - F = 0$, with E a diagonal matrix and F an M-matrix. Finally, in section 6 we use some numerical examples to illustrate the correctness of our theory and the feasibility of our methods.

* Received December 10, 2003; final revised September 6, 2003.
1) Subsidized by the Special Funds For Major State Basic Research Projects G1999032803 and the National Natural Science Foundation No. 10471146, China.

2. Basic Notations and Results

Given two matrices $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{m\times n}$, we write $A \geq B$ ($A > B$) if $a_{ij} \geq b_{ij}$ ($a_{ij} > b_{ij}$) hold for all i and j, and we call the matrix A positive (nonnegative), if $A > 0$ ($A \geq 0$).

Let $A \in \mathbb{R}^{n\times n}$. It is called a Z-matrix if all of its off-diagonal elements are nonpositive. Clearly, a Z-matrix $A \in \mathbb{R}^{n\times n}$ can be represented as $A = sI - B$, with $B \geq 0$. In particular, when $s > \rho(B)$, the spectral radius of the matrix B, A turns to a nonsingular M-matrix, and when $s = \rho(B)$, it turns to a singular M-matrix. We use $\lambda(A)$ to denote the spectrum of the matrix A, $\sigma_{\min}(A)$ the smallest singular value of A, and $\mathcal{R}(A)$ the range space spanned by the columns of the matrix A.

The open left (right) half plane is denoted by $\mathbb{C}_<$ ($\mathbb{C}_>$), and the closed left (right) half plane is denoted by $\mathbb{C}_\leq$ ($\mathbb{C}_\geq$), respectively. In addition, we use $\|\cdot\|$ to denote any consistent matrix norm on $\mathbb{C}^{n\times n}$ unless it is claimed explicitly. In particular, we use $\|\cdot\|_2$ and $\|\cdot\|_F$ to denote the spectral and the Frobenius norms of a matrix, respectively.

We recall that the separation of two matrices $B \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{m\times m}$ can be defined as follows. See [14].

$$\mathrm{sep}(B,C) := \inf\{\|PB - CP\| \mid B \in \mathbb{R}^{n\times n},\ C \in \mathbb{R}^{m\times m} \text{ and } P \in \mathbb{R}^{m\times n},\ \text{with } \|P\| = 1\}. \quad (3)$$

When the norm in (3) is specified to be the Frobenius norm, we denote the separation $\mathrm{sep}(B,C)$ by $\mathrm{sep}_F(B,C)$.

The following properties about an M-matrix can be found in [1].

Lemma 2.1. *[1] Given a Z-matrix $A \in \mathbb{R}^{n\times n}$. Then the following statements are equivalent:*

(a) A is a nonsingular M-matrix;

(b) $A^{-1} \geq 0$;

(c) $Av > 0$ holds for some vector $v > 0$;

(d) $\lambda(A) \subset \mathbb{C}_>$.

For the nonsymmetric ARE(1), from [2, 3] we know that the following results hold.

Lemma 2.2. *If the matrix K defined in (2) is a nonsingular M-matrix, then the ARE(1) has a minimal nonnegative solution S that satisfies that both matrices $D_C := D - CS$ and $A_C := A - SC$ are nonsingular M-matrices.*

Lemma 2.3. *If the matrix K defined in (2) is a nonsingular M-matrix, then the matrix H defined in (2) has n eigenvalues in $\mathbb{C}_>$ and m eigenvalues in $\mathbb{C}_<$.*

Lemma 2.4. *If the matrix K defined in (2) is a nonsingular M-matrix and S is a minimal nonnegative solution of the ARE(1), then*

$$\begin{pmatrix} I & 0 \\ S & I \end{pmatrix}\begin{pmatrix} D & C \\ -B & -A \end{pmatrix}\begin{pmatrix} I & 0 \\ -S & I \end{pmatrix} = \begin{pmatrix} D - CS & C \\ 0 & -(A - SC) \end{pmatrix}.$$

It then follows that the column space of the matrix

$$\begin{pmatrix} I \\ -S \end{pmatrix}$$

is the unique invariant subspace of the matrix H associated with its n eigenvalues in $\mathbb{C}_>$.

Conversely, let U be an orthogonal matrix such that

$$G := U^T H U$$

is a real Schur form of H, where all 1×1 or 2×2 diagonal blocks of H corresponding to the eigenvalues in $\mathbb{C}_>$ appear in the first n columns of G and all 1×1 or 2×2 diagonal blocks of H corresponding to the eigenvalues in $\mathbb{C}_<$ appear in the last m columns of G. If the matrix U is partitioned as

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \quad \text{with} \quad U_{11} \in \mathbb{R}^{n \times n},$$

then U_{11} is nonsingular and $S := -U_{21}U_{11}^{-1}$ is the minimal nonnegative solution of the ARE(1).

3. Perturbation Bounds

Throughout this section, we assume that the matrix K defined in (2) is a nonsingular M-matrix. Hence, from Lemma 2.2 we know that the nonsymmetric ARE(1) has a minimal nonnegative solution S and both matrices $D_C = D - CS$ and $A_C = A - SC$ are nonsingular M-matrices.

To investigate the variation in the minimal nonnegative solution S of the ARE(1) with respect to changes in the matrices A, B, C and D, we let $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$ and $\widetilde{D}$ be nearby matrices to A, B, C and D, respectively. Corresponding to (2) we define two matrices

$$\widetilde{H} = \begin{pmatrix} \widetilde{D} & \widetilde{C} \\ -\widetilde{B} & -\widetilde{A} \end{pmatrix}, \qquad \widetilde{K} = \begin{pmatrix} \widetilde{D} & -\widetilde{C} \\ -\widetilde{B} & \widetilde{A} \end{pmatrix}. \tag{4}$$

And we assume that the matrix $\widetilde{K}$ is also a nonsingular M-matrix. Let $\widetilde{S}$ be the minimal nonnegative solution of the nonsymmetric ARE

$$\widetilde{X}\widetilde{C}\widetilde{X} - \widetilde{X}\widetilde{D} - \widetilde{A}\widetilde{X} + \widetilde{B} = 0, \tag{5}$$

and define

$$\begin{aligned} \Delta A &= \widetilde{A} - A, \quad \Delta B = \widetilde{B} - B, \quad \Delta C = \widetilde{C} - C, \\ \Delta D &= \widetilde{D} - D, \quad \Delta S = \widetilde{S} - S, \quad \Delta H = \widetilde{H} - H. \end{aligned} \tag{6}$$

Then we are going to derive a bound on $\|\Delta S\|$, which should be reasonably sharp in certain sense.

To this end, we need the following results, which can be found in [11, 15, 16].

Lemma 3.1. *[11] Let $A, \widehat{A} \in \mathbb{C}^{n \times n}$ and $B, \widehat{B} \in \mathbb{C}^{m \times n}$. Assume that*

$$Z = \begin{pmatrix} A \\ B \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} \widehat{A} \\ \widehat{B} \end{pmatrix}$$

satisfy $Z^*Z = I$ and $W^*W = I$, where $(\cdot)^*$ denotes the conjugate transpose of the corresponding matrix. Let

$$\mathcal{X} = \mathcal{R}(Z), \qquad \mathcal{Y} = \mathcal{R}(W),$$

and

$$\Theta = \arccos(Z^*WW^*Z)^{\frac{1}{2}} \geq 0, \qquad d_F(\mathcal{X}, \mathcal{Y}) = \|\sin\Theta\|_F.$$

If the matrix A is nonsingular and $d_F(\mathcal{X},\mathcal{Y}) < \frac{1}{\sqrt{2}\sigma_{\min}(A)}$, then the matrix $\widehat{A}$ is also nonsingular and the matrices $X = BA^{-1}$ and $\widehat{X} = \widehat{B}\widehat{A}^{-1}$ satisfy

$$\|\widehat{X} - X\|_F \leq \frac{\sqrt{2}\|A^{-1}\|_2 d_F(\mathcal{X},\mathcal{Y})}{1-\sqrt{2}\|A^{-1}\|_2 d_F(\mathcal{X},\mathcal{Y})} \cdot \sqrt{1+\|X\|_2^2}.$$

Lemma 3.2. *[16] Let $S, E \in \mathbb{R}^{n\times n}$ with S being a positive stable matrix (i.e., $\lambda(S) \subset \mathbb{C}_>$). Define an operator $\mathbb{L} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ by*

$$\mathbb{L}(X) = S^T X + XS, \quad \forall X \in \mathbb{R}^{n\times n}$$

and the norm of its inverse by

$$\|\mathbb{L}^{-1}\| = \max_{X\in\mathbb{R}^{n\times n}\setminus\{0\}} \frac{\|\mathbb{L}^{-1}(X)\|}{\|X\|}.$$

If $\|\mathbb{L}^{-1}\|\|E\| < \frac{1}{2}$, then the matrix $S+E$ is positive stable.

Lemma 3.3. *[15] Let the matrices H, $\widetilde{H}$, K, $\widetilde{K}$ and ΔH be defined as in (2), (3), (4) and (6), respectively. Let $U = (U_1, U_2)$ be an orthogonal matrix with $\mathcal{R}(U_1)$ being an invariant subspace of H associated with n eigenvalues in $\mathbb{C}_>$. That is to say, U is an orthogonal matrix such that*

$$U^T H U = \begin{pmatrix} G_{11} & G_{12} \\ 0 & G_{22} \end{pmatrix},$$

where $G_{11} \in \mathbb{R}^{n\times n}$, $G_{22} \in \mathbb{R}^{m\times m}$, $\lambda(G_{11}) \subset \mathbb{C}_>$ and $\lambda(G_{22}) \subset \mathbb{C}_<$. Let $U^T \Delta H U$ be conformably partitioned as

$$U^T \ \Delta H \ U = \begin{pmatrix} \Delta G_{11} & \Delta G_{12} \\ \Delta G_{21} & \Delta G_{22} \end{pmatrix}.$$

Define

$$\delta = \mathrm{sep}_F(G_{11}, G_{22}) - (\|\Delta G_{11}\|_F + \|\Delta G_{22}\|_F) \quad \text{and} \quad \nu = \|\Delta G_{21}\|_F. \tag{7}$$

If

$$\delta > 4\|\Delta G_{21}\|_F(\|G_{12}\|_F + \|\Delta G_{12}\|_F),$$

then there exists a matrix $P \in \mathbb{R}^{m\times n}$, satisfying $\|P\|_F \leq \frac{2\nu}{\delta}$, such that the columns of the matrix

$$\widetilde{U}_1 := (U_1 + U_2 P)(I + P^T P)^{-\frac{1}{2}}$$

span an invariant subspace of $\widetilde{H}$.

We remark that, for the matrices U and P introduced in Lemma 3.3, if we define a matrix

$$\widetilde{U} := U \begin{pmatrix} I_n & -P^T \\ P & I_m \end{pmatrix} \begin{pmatrix} (I+P^TP)^{-\frac{1}{2}} & 0 \\ 0 & (I+PP^T)^{-\frac{1}{2}} \end{pmatrix}, \tag{8}$$

then it is easy to verify that $\widetilde{U}$ is an orthogonal matrix, and its first n columns span an invariant subspace of $\widetilde{H}$. This fact is precisely described in the following lemma.

Lemma 3.4. *[14] Under the conditions of Lemma 3.3, there exists an orthogonal matrix $\widetilde{U}$ such that*

$$\widetilde{U}^T \widetilde{H} \widetilde{U} = \begin{pmatrix} \widetilde{G}_{11} & \widetilde{G}_{12} \\ 0 & \widetilde{G}_{22} \end{pmatrix}. \tag{9}$$

Furthermore, if we conformably partition the matrices U and $\widetilde{U}$ into the block forms

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \quad \text{and} \quad \widetilde{U} = \begin{pmatrix} \widetilde{U}_{11} & \widetilde{U}_{12} \\ \widetilde{U}_{21} & \widetilde{U}_{22} \end{pmatrix}, \tag{10}$$

and denote

$$U_1 = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} \quad \text{and} \quad \widetilde{U}_1 = \begin{pmatrix} \widetilde{U}_{11} \\ \widetilde{U}_{21} \end{pmatrix},$$

then it holds that $d_F(\mathcal{X}, \mathcal{Y}) \leq \|P\|_F$, where $\mathcal{X} = \mathcal{R}(U_1)$ and $\mathcal{Y} = \mathcal{R}(\widetilde{U}_1)$.

It follows straightforwardly from (8) and (9) that

$$\widetilde{G}_{11} = (I + P^T P)^{\frac{1}{2}} [G_{11} + \Delta G_{11} + (G_{12} + \Delta G_{12}) P] (I + P^T P)^{-\frac{1}{2}}.$$

Hence,

$$\begin{aligned}
\|\widetilde{G}_{11} - G_{11}\|_F \;\; & \leq \;\; \|(I + P^T P)^{\frac{1}{2}} G_{11} (I + P^T P)^{-\frac{1}{2}} - G_{11}\|_F \\
& \quad + \|(I + P^T P)^{\frac{1}{2}}\|_2 [\|\Delta G_{11}\|_F + \|G_{12}\|_2 \|P\|_F + \|\Delta G_{12}\|_F \|P\|_F] \\
& \leq \;\; \|(I + P^T P)^{\frac{1}{2}} G_{11} (I + P^T P)^{-\frac{1}{2}} - G_{11} (I + P^T P)^{\frac{1}{2}} (I + P^T P)^{-\frac{1}{2}}\|_F \\
& \quad + \sqrt{1 + \|P\|_F^2} \cdot [(1 + \|P\|_F) \|\Delta H\|_F + \|G_{12}\|_2 \|P\|_F] \\
& \leq \;\; \|(I + P^T P)^{\frac{1}{2}} G_{11} - G_{11} (I + P^T P)^{\frac{1}{2}}\|_F \\
& \quad + \sqrt{1 + \|P\|_F^2} \cdot [(1 + \|P\|_F) \|\Delta H\|_F + \|G_{12}\|_2 \|P\|_F] \\
& \leq \;\; \|(I + P^T P)^{\frac{1}{2}} G_{11} - G_{11} + G_{11} - G_{11} (I + P^T P)^{\frac{1}{2}}\|_F \\
& \quad + \sqrt{1 + \|P\|_F^2} \cdot [(1 + \|P\|_F) \|\Delta H\|_F + \|G_{12}\|_2 \|P\|_F] \\
& \leq \;\; 2\|G_{11}\|_2 \|(I + P^T P)^{\frac{1}{2}} - I\|_F \\
& \quad + \sqrt{1 + \|P\|_F^2} \cdot [(1 + \|P\|_F) \|\Delta H\|_F + \|G_{12}\|_2 \|P\|_F] \\
& \leq \;\; 2\sqrt{2} \|G_{11}\|_2 \|P\|_F^2 + \sqrt{1 + \|P\|_F^2} \cdot [(1 + \|P\|_F) \|\Delta H\|_F + \|G_{12}\|_2 \|P\|_F] \\
& \leq \;\; 2\sqrt{2} \|G_{11}\|_2 \left(\frac{2\nu}{\delta}\right)^2 + \sqrt{1 + \left(\frac{2\nu}{\delta}\right)^2} \cdot \left[\left(1 + \frac{2\nu}{\delta}\right) \|\Delta H\|_F + \|G_{12}\|_2 \cdot \frac{2\nu}{\delta}\right] \\
& := \;\; \mu.
\end{aligned} \tag{11}$$

Since $\lambda(G_{11}) \subset \mathbb{C}_>$ implies that G_{11} is positive stable, we immediately know that the matrix

$$L_G := I \otimes G_{11}^T + G_{11}^T \otimes I \tag{12}$$

is also positive stable, where $\otimes$ denotes the Kronecker product. Therefore, the smallest singular value of L_G, say $\sigma_{\min}(L_G)$, is positive. Now, if we define an operator $\mathbb{L}_G : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ by

$$\mathbb{L}_G(X) = G_{11}^T X + X G_{11}, \qquad \forall X \in \mathbb{R}^{n \times n},$$

then $\mathbb{L}_G$ is invertible and the $\|\cdot\|_F$-norm of $\mathbb{L}_G^{-1}$ can be computed as follows:

$$\begin{aligned}
\|\mathbb{L}_G^{-1}\|_F &= \max_{X\in\mathbb{R}^{n\times n}\backslash\{0\}} \frac{\|\mathbb{L}_G^{-1}(X)\|_F}{\|X\|_F} = \max_{X\in\mathbb{R}^{n\times n},\ \|X\|_F=1} \|\mathbb{L}_G^{-1}(X)\|_F \\
&= \max_{x\in\mathbb{R}^{n^2},\ \|x\|=1} \|(I\otimes G_{11}^T + G_{11}^T\otimes I)^{-1}x\|_2 = \frac{1}{\sigma_{\min}(L_G)}. \qquad (13)
\end{aligned}$$

Based upon the above argument, we can easily obtain the perturbation bound of the minimal nonnegative solution of the nonsymmetric ARE(1) as follows.

Theorem 3.1. *Let all conditions of Lemma 3.3 be satisfied. If*

$$\frac{\mu}{\sigma_{\min}(L_G)} < \frac{1}{2} \quad \text{and} \quad \frac{2\|\Delta H\|_F}{\delta} < \frac{1}{\sqrt{2}\sigma_{\min}(U_{11})},$$

then

$$\|\widetilde{S} - S\|_F \leq \frac{2\sqrt{2}\|U_{11}^{-1}\|_2\|\Delta H\|_F}{\delta - 2\sqrt{2}\|U_{11}^{-1}\|_2\|\Delta H\|_F} \cdot \sqrt{1+\|S\|_2^2},$$

where the quantities δ and μ are defined in (7) and (11), and the matrices L_G and U_{11} are defined in (12) and (10), respectively.

Proof. Because

$$\frac{\mu}{\sigma_{\min}(L_G)} < \frac{1}{2},$$

from (11) and (13) we have

$$\|\mathbb{L}_G^{-1}\|_F\|\widetilde{G}_{11} - G_{11}\|_F < \frac{1}{2}.$$

According to Lemma 3.2, we know that the matrix $\widetilde{G}_{11}$ is positive stable, or in other words, $\lambda(\widetilde{G}_{11}) \subset \mathbb{C}_>$. Let $\widetilde{U}$ be partitioned as in (10). Then by making use of Lemmas 3.4 and 3.3 we obtain

$$d_F(\mathcal{X},\mathcal{Y}) \leq \|P\|_F \leq \frac{2\nu}{\delta} \leq \frac{2\|\Delta H\|_F}{\delta} < \frac{1}{\sqrt{2}\sigma_{\min}(U_{11})}.$$

Now, by Lemmas 3.1 and 2.4, we can immediately get that $\widetilde{U}_{11}$ is nonsingular, $\widetilde{S} = -\widetilde{U}_{21}\widetilde{U}_{11}^{-1}$ is the minimal nonnegative solution of the ARE(5), and

$$\|\widetilde{S} - S\|_F \leq \frac{2\sqrt{2}\|U_{11}^{-1}\|_2\|\Delta H\|_F}{\delta - 2\sqrt{2}\|U_{11}^{-1}\|_2\|\Delta H\|_F} \cdot \sqrt{1+\|S\|_2^2}$$

holds.

Theorem 3.1 clearly shows that the stability property of the minimal nonnegative solution of the ARE(1) is closely dependent on the quantity $\delta_F = \mathrm{sep}_F(G_{11}, G_{22})$. When δ_F is relatively large, the solution is insensitive to the perturbation of the matrix H, otherwise, it is very sensitive. Furthermore, $\delta_F = \|[G_{11}^T\otimes I - I\otimes G_{22}]^{-1}\|_2^{-1}$ is computable.

4. A Structured Condition Number

Consider the perturbed equation

$$\begin{aligned}&(S+\Delta S)(C+\Delta C)(S+\Delta S)\\&\quad -(S+\Delta S)(D+\Delta D)-(A+\Delta A)(S+\Delta S)+(B+\Delta B)=0,\end{aligned} \tag{14}$$

where

$$\Delta A\in\mathbb{R}^{m\times m},\quad \Delta B\in\mathbb{R}^{m\times n},\quad \Delta C\in\mathbb{R}^{n\times m},\quad \Delta D\in\mathbb{R}^{n\times n}\quad\text{and}\quad \Delta S\in\mathbb{R}^{m\times n}$$

are perturbing increments to the matrices A, B, C, D and S, respectively.

We are going to measure the perturbation ΔS normwisely by the quantity

$$\rho(\Delta A,\Delta B,\Delta C,\Delta D):=\left\|\left[\frac{\Delta A}{\alpha},\frac{\Delta B}{\beta},\frac{\Delta C}{\gamma},\frac{\Delta D}{\eta}\right]\right\|_F,$$

where α, β, γ and η are positive scalars. Similarly to Rice[12] we may define the condition number $\kappa(S)$ with respect to S by

$$\kappa(S)=\lim_{\delta\to\infty}\sup\left\{\frac{\|\Delta S\|_F}{\xi\delta}\mid K+\Delta K\quad\text{is an } M\text{-matrix}, \rho(\Delta A,\Delta B,\Delta C,\Delta D)\le\delta\right\}, \tag{15}$$

where ξ is a prescribed positive parameter, K is defined as in (2), and ΔK is defined by

$$\Delta K=\begin{pmatrix}\Delta D & -\Delta C\\ -\Delta B & \Delta A\end{pmatrix}.$$

Specializing

$$\xi=\alpha=\beta=\gamma=\eta=1$$

in (15) we get the absolute condition number, say $\kappa_{\text{abs}}(S)$, and taking

$$\xi=\|S\|_F,\quad \alpha=\|A\|_F,\quad \beta=\|B\|_F,\quad \gamma=\|C\|_F\quad\text{and}\quad \eta=\|D\|_F$$

in (15) we obtain the relative condition number, say $\kappa_{\text{rel}}(S)$.

Define $\kappa^*(S)$ to be the scalar

$$\kappa^*(S)=\lim_{\delta\to\infty}\sup\left\{\frac{\|\Delta S\|_F}{\xi\delta}\mid \rho(\Delta A,\Delta B,\Delta C,\Delta D)\le\delta\right\}. \tag{16}$$

Then we can analogously obtain the numbers $\kappa^*_{\text{abs}}(S)$ and $\kappa^*_{\text{rel}}(S)$.

From (15) and (16) we easily know that

$$\kappa(S)\le\kappa^*(S).$$

To derive an explicit expression for $\kappa^*(S)$, we expand (14) obtaining

$$(A-SC)\ \Delta S+\Delta S\ (D-CS)=-\Delta A\ S-S\ \Delta D+S\ \Delta C\ S+\Delta B+\mathcal{O}(\varrho^2). \tag{17}$$

Here, we have abbreviated $\rho(\Delta A,\Delta B,\Delta C,\Delta D)$ by ϱ. By using the vec operator, which stacks the columns of a matrix into one long vector, and considering the property

$$\text{vec}(AXB)=(B^T\otimes A)\cdot\text{vec}(X),$$

we can express (17) as the following:

$$\begin{aligned} P\cdot \mathrm{vec}(\Delta S) &= -(S^T\otimes I)\cdot \mathrm{vec}(\Delta A)-(I\otimes S)\cdot \mathrm{vec}(\Delta D)\\ &\quad +(S^T\otimes S)\cdot \mathrm{vec}(\Delta C)+\mathrm{vec}(\Delta B)+\mathcal{O}(\varrho^2)\\ &= [-\alpha(S^T\otimes I),\beta I_{n^2},\gamma(S^T\otimes S),-\eta(I\otimes S)]\vec{\Delta}_{\alpha,\beta,\gamma,\eta}+\mathcal{O}(\varrho^2). \end{aligned} \tag{18}$$

Here, we have adopted the notations

$$\begin{aligned} P &= I\otimes(A-SC)+(D-CS)^T\otimes I,\\ \vec{\Delta}_{\alpha,\beta,\gamma,\eta} &= \begin{bmatrix} \frac{\mathrm{vec}(\Delta A)}{\alpha}\\ \frac{\mathrm{vec}(\Delta B)}{\beta}\\ \frac{\mathrm{vec}(\Delta C)}{\gamma}\\ \frac{\mathrm{vec}(\Delta D)}{\eta}\end{bmatrix}. \end{aligned}$$

Noticing that P is nonsingular. Hence, by premultiplying by P^{-1} and then taking the 2-norm on both sides of the equality (18), we obtain

$$\begin{aligned} \kappa^*(S) &= \frac{1}{\xi}\cdot\max_{\varrho\neq 0}\frac{\|\Delta S\|_F}{\varrho}\\ &= \frac{1}{\xi}\cdot\max_{\varrho\neq 0}\frac{\|P^{-1}[-\alpha(S^T\otimes I),\beta I_{n^2},\gamma(S^T\otimes S),-\eta(I\otimes S)]\vec{\Delta}_{\alpha,\beta,\gamma,\eta}\|_2}{\|\vec{\Delta}_{\alpha,\beta,\gamma,\eta}\|_2}\\ &= \frac{1}{\xi}\|P^{-1}[-\alpha(S^T\otimes I),\beta I_{n^2},\gamma(S^T\otimes S),-\eta(I\otimes S)]\|_2, \end{aligned}$$

where we have used the fact $\|\mathrm{vec}(X)\|_2=\|X\|_F$. Now, by letting

$$\xi=\|S\|,\quad \alpha=\|A\|,\quad \beta=\|B\|,\quad \gamma=\|C\|,\quad \eta=\|D\|,$$

based on (18) we can get the following estimate about the relative perturbation with respect to the solution S:

$$\frac{\|\Delta S\|}{\|S\|}\leq \kappa^*_{\mathrm{rel}}(S)\varrho+\mathcal{O}(\varrho^2).$$

The following example shows the tightness of the above-derived perturbation bounds.

Example 4.1. Consider the nonsymmetric ARE(1), for which

$$A=D=\mathrm{Tridiag}(-I,T,-I)\in\mathbb{R}^{n\times n}$$

are block tridiagonal matrices,

$$B=C=\frac{1}{50}\cdot\mathrm{tridiag}(1,2,1)\in\mathbb{R}^{n\times n}$$

are tridiagonal matrices, where

$$T=\mathrm{tridiag}(-1,4+ch^2,-1)\in\mathbb{R}^{m\times m}$$

is a tridiagonal matrix, I the m-by-m identity matrix , $n=m^2$, and $h=\frac{1}{m+1}$.

Let A, B, C and D be perturbed to $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$ and $\widetilde{D}$, respectively, where

$$\widetilde{A}=A+10^{-j}C,\quad \widetilde{B}=B,\quad \widetilde{C}=C,\quad \widetilde{D}=D,$$

and j is an integer. Assume that S and $\widetilde{S}$ are the minimal nonnegative solutions of the ARE(1) and the ARE(5), respectively. The numerical results in Table 1 shows that the relative perturbation bounds and the condition numbers derived in this section are quite feasible.

Table 1: Relative perturbation bounds and condition numbers for Example 4.1

j	$\frac{\|\widetilde{S}-S\|_F}{\|S\|_F}$	$\kappa^*_{\rm rel}(S)\rho$	$\kappa^*_{\rm rel}(S)$
-2	4.1987e-005	9.0991e-005	1.8739
-4	4.1987e-007	9.0991e-007	1.8739
-6	4.1987e-009	9.0991e-009	1.8739
-8	4.1793e-011	9.0991e-011	1.8739
-10	4.0789e-013	9.0991e-013	1.8739

5. Numerical Methods

Let H and K be the matrices defined in (2). If K is a nonsingular M-matrix, then by Lemma 2.3 we know that the matrix H has n eigenvalues in $\mathbb{C}_>$ and m eigenvalues in $\mathbb{C}_<$. Based on Lemma 2.4, we can compute the minimal nonnegative solution of the nonsymmetric ARE(1) by the Schur decomposition of the matrix H.

In a different way, in this section we will present a new method for computing the minimal nonnegative solution of the nonsymmetric ARE(1) through the matrix sign function method, and we will use this new method to find the nonsingular M-matrix solution of the quadratic matrix equation $X^2 - EX - F = 0$, where E is a diagonal matrix and F is an M-matrix.

5.1 The matrix sign function method

For a complex $z \in \mathbb{C}$ satisfying $\mathrm{Re}(z) \neq 0$, we define its sign by

$$\mathrm{sign}(z) = \begin{cases} 1, & \text{if} \quad \mathrm{Re}(z) > 0, \\ -1, & \text{if} \quad \mathrm{Re}(z) < 0. \end{cases}$$

Similarly, for a matrix $Z \in \mathbb{R}^{n\times n}$ satisfying $\mathrm{Re}(\lambda) \neq 0$, $\forall \lambda \in \lambda(Z)$, let T be a nonsingular matrix such that

$$Z = T \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} T^{-1}$$

is the Jordan decomposition of Z, where the Jordan blocks J_1 corresponding to eigenvalues in $\mathbb{C}_>$ and J_2 corresponding to those in $\mathbb{C}_<$, respectively, of the matrix Z. Then we can define the sign function of the matrix Z by

$$\mathrm{sign}(Z) = T \begin{pmatrix} I_1 & 0 \\ 0 & -I_2 \end{pmatrix} T^{-1},$$

where I_1 and I_2 are identity matrices of the same sizes as J_1 and J_2, respectively.

The following theorem shows how to use this matrix sign function to find the minimal nonnegative solution of the nonsymmetric ARE(1) when the matrix K defined in (2) is a nonsingular M-matrix.

Theorem 5.1. *Let*

$$W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} := sign(H),$$

and

$$\overline{W}_1 = \begin{pmatrix} W_{11} - I_n \\ W_{21} \end{pmatrix}, \quad \overline{W}_2 = \begin{pmatrix} W_{12} \\ W_{22} - I_m \end{pmatrix}.$$

Then $\overline{W}_2$ is of full column rank and the minimal nonnegative solution S of the ARE(1) is the unique least-squares solution of the matrix equation $\overline{W}_2 S = \overline{W}_1$.

Proof. Denote by

$$T := \begin{pmatrix} I_n & 0 \\ -S & I_m \end{pmatrix} \begin{pmatrix} I_n & Y \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} I_n & Y \\ -S & I_m - SY \end{pmatrix},$$

where S is the minimal nonnegative solution of the nonsymmetric ARE(1) and Y is the solution of the Sylvester equation

$$D_C Y + Y A_C + C = 0,$$

with

$$D_C = D - CS \quad \text{and} \quad A_C = A - SC.$$

See Lemma 2.2. Then we easily have

$$T^{-1} H T = \begin{pmatrix} D_C & 0 \\ 0 & -A_C \end{pmatrix}.$$

It then follows that

$$\mathrm{sign}(H) = T \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix} T^{-1} \equiv \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} T = T \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix},$$

or equivalently,

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} T = T \begin{pmatrix} 0 & 0 \\ 0 & -2I_m \end{pmatrix} + T \begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix}.$$

This clearly shows that

$$\begin{pmatrix} W_{11} - I_n & W_{12} \\ W_{21} & W_{22} - I_m \end{pmatrix} T = T \begin{pmatrix} 0 & 0 \\ 0 & -2I_m \end{pmatrix}.$$

By substituting the actual expression of T into this equality, we get

$$\begin{pmatrix} W_{11} - I_n & W_{12} \\ W_{21} & W_{22} - I_m \end{pmatrix} \begin{pmatrix} I_n & Y \\ -S & I_m - SY \end{pmatrix} = \begin{pmatrix} I_n & Y \\ -S & I_m - SY \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2I_m \end{pmatrix}. \tag{19}$$

Evidently, the first n columns of the matrix equation (19) satisfy

$$\begin{pmatrix} W_{11} - I_n & W_{12} \\ W_{21} & W_{22} - I_m \end{pmatrix} \begin{pmatrix} I_n \\ -S \end{pmatrix} = 0,$$

which implies

$$\begin{pmatrix} W_{12} \\ W_{22} - I_m \end{pmatrix} S = \begin{pmatrix} W_{11} - I_n \\ W_{21} \end{pmatrix},$$

or

$$\overline{W}_2 S = \overline{W}_1. \tag{20}$$

We can further demonstrate that the matrix $\overline{W}_2$ is of full column rank, and hence, the matrix equation (20) has a unique least-squares solution. In fact, from (19) we have

$$\begin{pmatrix} W_{11} - I_n & W_{12} \\ W_{21} & W_{22} - I_m \end{pmatrix} = \begin{pmatrix} -2YS & -2Y \\ -2S + 2SYS & -2 + 2SY \end{pmatrix}.$$

This implies that

$$\begin{pmatrix} W_{12} \\ W_{22} - I_m \end{pmatrix} = -2 \begin{pmatrix} Y \\ I_m - SY \end{pmatrix},$$

and hence, $\overline{W}_2$ is of full column rank.

The matrix sign function method can be algorithmically described as follows.

Algorithm 5.1. *The algorithm consists of the following three steps:*

(i) Let

$$H = \begin{pmatrix} D & C \\ -B & -A \end{pmatrix}.$$

Compute the matrix sign function

$$W = \mathrm{sign}(H) \equiv \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix};$$

(ii) Let

$$\overline{W}_1 = \begin{pmatrix} W_{11} - I_n \\ W_{21} \end{pmatrix} \quad \textit{and} \quad \overline{W}_2 = \begin{pmatrix} W_{12} \\ W_{22} - I_n \end{pmatrix};$$

(iii) Solve the linear matrix equation $\overline{W}_2 S = \overline{W}_1$.

In Algorithm 5.1 we need to compute $W = \mathrm{sign}(H)$. There are several methods to do this. Below we only introduce two of them which is particularly effective to solve our problem.

First, by adopting the expression $\mathrm{sign}(H) = H(H^2)^{-\frac{1}{2}}$ introduced in Higham [4], we can obtain $\mathrm{sign}(H)$ through solving the linear matrix equation $\mathrm{sign}(H)(H^2)^{\frac{1}{2}} = H$. Noticing that H^2 has no nonpositive real eigenvalues, we know that the matrix $(H^2)^{\frac{1}{2}}$ is well defined and can be computed through the Schur decomposition method[5].

Second, we can obtain $\mathrm{sign}(H)$ through the Newton iteration

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}), \qquad k = 0, 1, 2, \ldots,$$

where $X_0 = H$ is the initial guess and $\mathrm{sign}(H) = \lim_{k\to\infty} X_k$. Kenney and Laub have proved in [7] that this Newton iteration method is quadratically convergent and is also stable.

From the above investigation, we see that Algorithm 5.1 is stable with low cost. However, the *Schur decomposition method* proposed in [3] needs ordering according to the matrix eigenvalues which may lead to large cost, and the matrix U_{11} involved in it may be nearly singular even the matrix H is strongly nonsingular. See Lemma 2.4 for the definitions of the matrices U_{11} and H. This observations will be further validated by Examples 6.1 and 6.2.

5.2 Application to the quadratic matrix equation

Consider the *quadratic matrix equation* **(QME)**

$$X^2 - EX - F = 0, \tag{21}$$

where E is an n-by-n diagonal matrix and F is an n-by-n nonsingular M-matrix.

From [3] we know that the QME(21) has the following property.

Lemma 5.1. *[3] Let $E, F \in \mathbb{R}^{n\times n}$, with E being a diagonal matrix and F a nonsingular M-matrix. Then the QME(21) has a unique nonsingular M-matrix solution.*

In [3], the Schur decomposition method (see Lemma 2.4) was used to compute the M-matrix solution of the QME(21). In a quite different way, in the following we will use the matrix sign function method described in section 5.1 to find the M-matrix solution of the QME(21).

To this end, for a given parameter $\alpha > 0$, we let $Y = \alpha I - X$ and rewrite (21) as

$$Y^2 - Y(\alpha I) - (\alpha I - E)Y + (\alpha^2 I - \alpha E - F) = 0. \tag{22}$$

Hence, the QME(21) is equivalent to the nonsymmetric ARE(1) with the following specific choices of the involved matrices:

$$A = \alpha I - E, \quad B = \alpha^2 I - \alpha E - F, \quad C = I \quad \text{and} \quad D = \alpha I.$$

Now, the matrices H and K defined in (2) possess the specific forms

$$H := H_\alpha = \begin{pmatrix} \alpha I & I \\ -\alpha^2 I + \alpha E + F & -(\alpha I - E) \end{pmatrix}$$

and

$$K := K_\alpha = \begin{pmatrix} \alpha I & -I \\ -\alpha^2 I + \alpha E + F & \alpha I - E \end{pmatrix},$$

respectively.

Let e_i and f_i be the i-th diagonal elements of the matrices E and F, respectively. Then for $\alpha \in [\alpha_0, +\infty)$, with

$$\alpha_0 = \max_{1\le i\le n} \frac{e_i + \sqrt{e_i^2 + 4f_i}}{2} > 0,$$

we have

$$\alpha > 0 \quad \text{and} \quad \alpha^2 I - \alpha E - F \ge 0.$$

Moreover, from [3] we know that K_α is a nonsingular M-matrix. Therefore, by Lemma 2.2 the nonsymmetric ARE(22) has a minimal nonnegative solution S_α which satisfies that both matrices $\alpha I - S_\alpha$ and $\alpha I - E - S_\alpha$ are nonsingular M-matrices. And by Lemma 2.3 the matrix H_α has n eigenvalues in $\mathbb{C}_>$ and n eigenvalues in $\mathbb{C}_<$. Clearly, $\alpha I - S_\alpha$ is the M-matrix solution of the QME(21).

Denote by

$$W = \begin{pmatrix} 0 & I \\ F & E \end{pmatrix}.$$

Because

$$\begin{pmatrix} I & 0 \\ -\alpha I & I \end{pmatrix}^{-1} H_\alpha \begin{pmatrix} I & 0 \\ -\alpha I & I \end{pmatrix} = W,$$

we see that the matrices H_α and W have the same eigenvalues. Therefore, if X is the M-matrix solution of the QME(21), then

$$W \begin{pmatrix} I & 0 \\ X & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ X & I \end{pmatrix} \begin{pmatrix} X & I \\ 0 & E - X \end{pmatrix},$$

which implies that the column space of the matrix

$$\begin{pmatrix} I \\ X \end{pmatrix}$$

is the unique invariant subspace of W associated with the n eigenvalues of the matrix H_α in $\mathbb{C}_>$.

Now, through a similar discussion to Theorem 5.1, we can obtain the following theorem.

Theorem 5.2. *Let*

$$W = \begin{pmatrix} 0 & I \\ F & E \end{pmatrix}$$

and denote

$$\operatorname{sign}(W) = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

Then the M-matrix solution X of the QME(21) is the unique least-squares solution of the matrix equation

$$\begin{pmatrix} V_{12} \\ V_{22} - I_n \end{pmatrix} X = - \begin{pmatrix} V_{11} - I_n \\ V_{21} \end{pmatrix},$$

where the matrix

$$\begin{pmatrix} V_{12} \\ V_{22} - I_n \end{pmatrix}$$

is of full column rank.

We remark that the above-mentioned matrix sign function method can also be used to compute the following Wiener-Hopf factorization of Markov chains[13, 2]:

$$\begin{pmatrix} A & B \\ -C & -D \end{pmatrix} \begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix} = \begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix} \begin{pmatrix} \widetilde{Q} & 0 \\ 0 & \widehat{Q} \end{pmatrix}, \tag{23}$$

where A and D are square matrices such that

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is a nonsingular Q-matrix associated with an irreducible continuous-time finite Markov chain, and $\widetilde{Q}$ and $\widehat{Q}$ are Q-matrices, too.

As a matter of fact, noticing that (23) is equivalent to

$$\begin{aligned} &\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} A & B \\ -C & -D \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \\ = &\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} \widetilde{Q} & 0 \\ 0 & -\widehat{Q} \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \end{aligned}$$

or

$$\begin{pmatrix} -D & -C \\ B & A \end{pmatrix}\begin{pmatrix} I & \Pi_1 \\ \Pi_2 & I \end{pmatrix} = \begin{pmatrix} I & \Pi_1 \\ \Pi_2 & I \end{pmatrix}\begin{pmatrix} -\widetilde{Q} & 0 \\ 0 & \widehat{Q} \end{pmatrix},$$

after a similar discussion to Theorem 5.1, we can demonstrate that Π_1 and Π_2 are the unique least-squares solutions of the matrix equations

$$\begin{pmatrix} Z_{11} + I_n \\ Z_{21} \end{pmatrix}\Pi_1 = -\begin{pmatrix} Z_{12} \\ Z_{22} + I_m \end{pmatrix}$$

and

$$\begin{pmatrix} Z_{12} \\ Z_{22} - I_m \end{pmatrix}\Pi_2 = -\begin{pmatrix} Z_{11} - I_n \\ Z_{21} \end{pmatrix},$$

respectively, provided

$$\begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} = \operatorname{sign}\begin{pmatrix} -D & -C \\ B & A \end{pmatrix}.$$

6. Numerical Examples

We now compare computational accuracy and effectiveness of the *matrix sign function* (**MSF**) method and the *Schur decomposition* (**SD**) method [3] in terms of the *residual error* (**RES**) and the *CPU timing* (**CPU**) of the ARE(1) and the QME(21). All results are produced by using MATLAB 6.5.

Example 6.1. We consider the nonsymmetric ARE(1) generated by the following process: Firstly, generate a 100×100 random matrix with no zero elements and save it as R; Secondly, let $W = \operatorname{diag}(R \cdot e) - R$, where $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$; And finally, for $\alpha = 1, 5, 9$, define

$$K := \begin{pmatrix} D & -C \\ -B & A \end{pmatrix} \equiv \alpha I + W.$$

Then K is a nonsingular M-matrix, and the nonsymmetric ARE(1) with the matrices A, B, C and D described above has a unique minimal nonnegative solution.

From the results in Table 2, we clearly see that the matrix sign function method is more accurate and effective than the Schur decomposition method, as the former always yields smaller residual error and computing time than the latter for all of our tested cases.

A matrix is called a Q-matrix if it has nonnegative off-diagonal elements and nonpositive row sums.

Table 2: Numerical Results for Example 6.1

Method		$\alpha = 1$	$\alpha = 5$	$\alpha = 9$
SD	RES	0.16e-11	0.17e-11	0.24e-11
	CPU	0.359	0.359	0.344
MSF	RES	0.73e-12	0.75e-12	0.66e-12
	CPU	0.297	0.313	0.312

Example 6.2. We consider the QME(21) generated by the matrices

$$E = \begin{bmatrix} -I_{\frac{n}{2}} & 0 \\ 0 & 3I_{\frac{n}{2}} \end{bmatrix} \in \mathbb{R}^{n\times n} \quad \text{and} \quad F = \begin{bmatrix} 2 & -1 & & & \\ & 2 & -1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ -1 & & & & 2 \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

As F is a nonsingular M-matrix, the QME(21) has a nonsingular M-matrix solution.

Table 3: Numerical Results for Example 6.2

Method		$n = 64$	$n = 96$	$n = 128$
SD	RES	0.50e-12	0.83e-12	0.11e-11
	CPU	0.672	4.344	15.016
MSF	RES	0.27e-12	0.38e-12	0.48e-12
	CPU	0.593	1.375	2.875

Again, from the results in Table 3, we see that the matrix sign function method is more accurate and effective than the Schur decomposition method, as the former always yields smaller residual error and computing time than the latter for all of our tested cases.

References

[1] A. Berman and R.J. Plemmons, Nonnegative Matrices in Mathematical Sciences, *SIAM*, Philadelphia, PA, 1994.

[2] C.-H. Guo, Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices, *SIAM J. Matrix Anal. Appl.*, **23** (2001), 225-242.

[3] C.-H. Guo, On a quadratic matrix equation associated with an M-matrix, *IMA J. Numer. Anal.*, **23** (2003), 11-27.

[4] N.J. Higham, The matrix sign decomposition and its relation to the polar decomposition, *Linear Algebra Appl.*, **212/213** (1994), 3-20.

[5] N.J. Higham, Computing real square roots of a real matrix, *Linear Algebra Appl.*, **88/89** (1987), 405-430.

[6] J. Juang and W.-W. Lin, Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices, *SIAM J. Matrix Anal. Appl.*, **20** (1998), 228-243.

[7] C. Kenney and A.J. Laub, On scaling Newton's method for polar decomposition and the matrix sign function, *SIAM J. Matrix Anal. Appl.*, **13** (1992), 688-706.

[8] A.-P. Liao, A class of inverse problems of matrix equation $AX = B$ and its numerical solution on the linear manifold, *Math. Numer. Sinica*, **20** (1998), 371-376.

[9] A.-P. Liao, On the least squares problem of a matrix equation, *J. Comput. Math.*, **17** (1999), 589-594.

[10] X.-G. Liu, The condition numbers of stabilizing solutions to algebraic Riccati equations, *Math. Numer. Sinica*, **23** (2001), 71-80.

[11] X.-G. Liu and X.-X. Guo, Sensitivity analysis of the discrete-time algebraic Riccati equation, *Math. Numer. Sinica*, **21** (1999), 163-170.

[12] J.R. Rice, A theory of condition, *SIAM J. Numer. Anal.*, **3** (1966), 287-310.

[13] L.C.G. Rogers, Fluid models in queuing theory and Wiener-Hopf factorization of Markov chains, *Ann. Appl. Probab.*, **4** (1994), 390-413.

[14] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.*, **15** (1973), 727-764.

[15] J.-G. Sun, Matrix Perturbation Theory, *Academic Press*, Beijing, 1987.

[16] J.-G. Sun, Perturbation Analysis of Algebraic Riccati Equations, *Technical Report UMINF 02.03*, Department of Computing Science, Umea Univ., 2002.

[17] S.-F. Xu, Sensitivity analysis of the algebraic Riccati equations, *Numer. Math.*, **75** (1996), 121-134.

[18] S.-F. Xu, Lower bound estimation for the separation of two matrices, *Linear Algebra Appl.*, **262** (1997), 67-82.