TAIWANESE JOURNAL OF MATHEMATICS
Vol. 10, No. 5, pp. 1329-1344, September 2006
This paper is available online at http://www.math.nthu.edu.tw/tjm/

PERTURBATION ANALYSIS OF A NONLINEAR MATRIX EQUATION

Shufang Xu and Mingsong Cheng

Abstract. Consider the nonlinear matrix equation $X + A^*X^{-2}A = I$, where A is an $n \times n$ complex matrix, I the identity matrix and A^* the conjugate transpose of the matrix A. In this paper a perturbation bound for a class of special solutions of this matrix equation is derived, and an explicit expression of its condition number is obtained. The results are illustrated by using some numerical examples.

1. INTRODUCTION

Consider the nonlinear matrix equation

$$X + A^*X^{-2}A = I, \tag{1.1}$$

where $A \in \mathbf{C}^{n\times n}$ and I denotes the identity matrix. Here $\mathbf{C}^{n\times n}$ denotes the set of all $n \times n$ complex matrices, A^* the conjugate transpose of a matrix A. It is often required to find the Hermitian positive definite (h.p.d.) solutions of the matrix Equation (1.1).

Ivanov et al. [1, 4, 5] gave some sufficient conditions for the existence of the Hermitian positive definite solutions to the matrix Equation (1.1) and proposed some iterative methods for solving it. Liu and Gao [6] consider the more general matrix equation

$$X^s + A^TX^{-t}A = I, \tag{1.2}$$

where s and t are natural numbers. They derived some sufficient conditions for the existence of a positive definite solution based on the fixed-point theory, and proposed three iterative methods for computing the positive definite solutions. See also [2, 7. 9-11].

Received November 16, 2004.
Communicated by Wen-Wei Lin.
2000 *Mathematics Subject Classification*: 15A24, 65F10, 65H05.
Key words and phrases: Nonlinear matrix equation, Perturbation bound, Condition number.

In this paper a perturbation bound for a class of special solutions of the matrix Equation (1.1) is derived, and an explicit expression of its condition number is obtained. The results are illustrated by using some numerical examples.

We start with some notations which we shall use throughout this paper. We use $\mathbf{H}^{n\times n}$ to denote the set of all $n \times n$ Hermitian matrices. A^T, $\bar{A}$, and A^* denote the transpose, the conjugate, and the conjugate transpose of a matrix A, respectively. The symbols $\|\cdot\|$ denotes the matrix spectral norm, and $\|\cdot\|_F$ the Frobenius norm. For $A = [a_1, \dots, a_n] = [a_{ij}] \in \mathbf{C}^{n\times n}$ and $B \in \mathbf{C}^{m\times m}$, we use $A \otimes B = [a_{ij}B]$ to denotes the Kronecker product of A and B, and use $\mathrm{vec}(A)$ to denote the vector defined by $\mathrm{vec}(A) = (a_1^T, \dots, a_n^T)^T$. For Hermitian mtrices P and Q, we write $P > Q$ ($P \geq Q$) if $P - Q$ is Hermitian positive definite (semi-definite). It is well known that $P \geq Q > 0$ implies $P^{-1} \leq Q^{-1}$ and $P^{\frac{1}{2}} \geq Q^{\frac{1}{2}}$.

The paper is organized as follows. In Section 2, we give the definition of the maximal solution to the matrix Equation (1.1), and present some of its properties. In Section 3, we derive a perturbation bound for a class of special maximal solution to the matrix Equation (1.1). An explicit expression of the condition number for the maximal solution is given in Section 4. Finally, some illustration numerical examples are given in Section 5.

2. Maximal Solution and Its Properties

Let

$$\mathcal{H} = \{\, X \;:\; X \text{ is a h.p.d. solution of the matrix Equation (1.1)} \,\}. \tag{2.1}$$

If $X \in \mathcal{H}$ satisfies that for any $Y \in \mathcal{H}$, $Y \geq X$ implies $X = Y$, then X is called a *maximal solution* of the matrix Equation (1.1).

Theorem 2.1. *There is at most one $X \in \mathcal{H}$ such that $X > \frac{2}{3}I$. Therefore, If $X \in \mathcal{H}$ satisfies that $X > \frac{2}{3}I$, then X is a maximal solution of the matrix equation.*

To prove this theorem, we need the following lemma.

Lemma 2.1. [4]. *The matrix Equation (1.1) has a h.p.d. solution X if and only if there exist unitary matrices U and V and diagonal matrices $\Gamma > 0$ and $\Sigma \geq 0$ such that $\Gamma + \Sigma^2 = I$ and $A = V\Gamma U\Sigma V^*$. In this case, $X = V\Gamma V^*$ is a h.p.d. solution.*

Proof of Theorem 2.1. Suppose $X_1,\ X_2 \in \mathcal{H}$ with $X_1,\ X_2 > \frac{2}{3}I$, and let $X_1 = V_1\Gamma_1V_1^*$ and $X_2 = V_2\Gamma_2V_2^*$ be the spectral decompositions of X_1 and X_2, respectively. By Lemma 2.1, there exist unitary matrices U_1, U_2 and diagonal matrices $\Sigma_1 \geq 0$, $\Sigma_2 \geq 0$ such that $\Gamma_1 + \Sigma_1^2 = I$, $\Gamma_2 + \Sigma_2^2 = I$, and $A = V_1\Gamma_1 U_1 \Sigma_1 V_1^* = V_2 \Gamma_2 U_2 \Sigma_2 V_2^*$. It follows from $\Gamma_1 + \Sigma_1^2 = I$ and $\Gamma_1 = V_1^* X_1 V_1 > \frac{2}{3}I$ that $\|\Sigma_1\| < \frac{\sqrt{3}}{3}$, and hence

$$\|X_1^{-1}A\| = \|V_1\Gamma_1^{-1}V_1^*V_1\Gamma_1U_1\Sigma_1V_1^*\| = \|\Sigma_1\| < \frac{\sqrt{3}}{3},$$
$$\|X_1^{-2}A\| = \|V_1\Gamma_1^{-2}V_1^*V_1\Gamma_1U_1\Sigma_1V_1^*\| \leq \|\Gamma_1^{-1}\|\|\Sigma_1\| < \frac{\sqrt{3}}{2}.$$

Similarly, we have $\|X_2^{-1}A\| < \frac{\sqrt{3}}{3}$ and $\|X_2^{-2}A\| < \frac{\sqrt{3}}{2}$.

Let $\Delta X = X_2 - X_1$, then ΔX satisfies

$$\Delta X - A^* X_1^{-2} \Delta X X_2^{-1} A - A^* X_1^{-1} \Delta X X_2^{-2} A = 0. \tag{2.2}$$

Denote $\mathbf{F}(W) = W - A^*X_1^{-2}WX_2^{-1}A - A^*X_1^{-1}WX_2^{-2}A$, then

$$\|\mathbf{F}(W)\| \geq \|W\|\big(1 - \|X_1^{-2}A\|\|X_2^{-1}A\| - \|X_1^{-1}A\|\|X_2^{-2}A\|\big) = \rho\|W\|,$$

where $\rho = 1 - \|X_1^{-2}A\|\|X_2^{-1}A\| - \|X_1^{-1}A\|\|X_2^{-2}A\| > 0$. This shows that $\mathbf{F}$ is invertible, and so $\mathbf{F}(\Delta X) = 0$ has a unique solution $\Delta X = 0$, that is, $X_1 = X_2$. This completes the proof. ∎

Theorem 2.2. *If $X \in \mathcal{H}$ satisfies $\|X^{-1}\|^3\|A\|^2 < \frac{1}{2}$, then $X > \frac{2}{3}I$, and so X is a maximal solution.*

Proof. Suppose that $X > \frac{2}{3}I$ is not true, then the smallest eigenvalue λ_n of X must satisfy $\lambda_n \leq \frac{2}{3}$. Hence we have $\|X^{-1} - I\| \geq \frac{1}{2}$. On the other hand, the assumption $\|X^{-1}\|^3\|A\|^2 < \frac{1}{2}$ implies that

$$\begin{aligned}\|X^{-1} - I\| &= \|X^{-1}(I - X)\| \leq \|X^{-1}\|\,\|I - X\| \\ &= \|X^{-1}\|\,\|A^*X^{-2}A\| \leq \|X^{-1}\|^3\,\|A\|^2 \\ &< \frac{1}{2}.\end{aligned}$$

This is a contradiction, hence we have $X > \frac{2}{3}I$. Thus, By Theorem 2.1, X is a maximal solution of the matrix Equation (1.1). ∎

Theorem 2.3. *If $\|A\| < \frac{2\sqrt{3}}{9}$, then there exists $X_L \in \mathcal{H}$ such that $X_L > \frac{2}{3}I$. Moreover, we have*

$$\|X_L^{-1}\| \leq 1 + \eta, \tag{2.3}$$

and for any other solution X (including non-Hermitian ones) of the matrix Equation (1.1),

$$\|X^{-1}\| > 2 - \eta. \tag{2.4}$$

Here $0 < \eta < \frac{1}{2}$ *satisfies*

$$\|A\|^2(1+\eta)^3 = \eta. \tag{2.5}$$

Proof. It is easy to verify that X is a solution to the matrix Equation (1.1) if and only if $Y = X^{-1}$ satisfies

$$Y = A^*Y^2AY + I. \tag{2.6}$$

Now define matrix sequence $\{Y_k\}$ by

$$Y_k = I + A^*Y_{k-1}^2AY_{k-1}, \tag{2.7}$$

where $Y_0 = 0$. By induction we can prove that

$$\|Y_k\| \leq 1 + \eta_k, \quad k = 1, 2, \ldots, \tag{2.8}$$

where

$$\eta_1 = 0, \quad \eta_k = \|A\|^2(1+\eta_{k-1})^3. \tag{2.9}$$

From (2.9) and the assumption $\|A\| < \frac{2\sqrt{3}}{9}$ we can easily derive that

$$\frac{1}{2} > \eta_{k+1} > \eta_k \geq 0,$$

and hence there exists a positive number η with $0 < \eta \leq \frac{1}{2}$ such that $\eta = \lim_{k\to\infty} \eta_k$. Combining this with (2.9) gives rise to that

$$\|A\|^2(1+\eta)^3 = \eta,$$

which, together with $\|A\| < \frac{2\sqrt{3}}{9}$, implies that $\eta < \frac{1}{2}$. Consequently, from (2.8), we get

$$\|Y_k\| \leq 1 + \eta_k \leq 1 + \eta,$$

which implies that

$$\begin{aligned}
\|Y_{k+1} - Y_k\| &= \|A^*Y_k^2AY_k - A^*Y_{k-1}^2AY_{k-1}\| \\
&\leq \|A\|\|Y_k^2AY_k - Y_k^2AY_{k-1} + Y_k^2AY_{k-1} - Y_{k-1}^2AY_{k-1}\| \\
&\leq \|A\|\big(\|Y_k^2A\|\|Y_k - Y_{k-1}\| + \|Y_k^2 - Y_{k-1}^2\|\|AY_{k-1}\|\big) \\
&\leq \|A\|^2\big(\|Y_k\|^2 + \|Y_{k-1}\|\|Y_k\| + \|Y_{k-1}\|^2\big)\|Y_k - Y_{k-1}\| \\
&\leq 3(1+\eta)^2\|A\|^2\|Y_k - Y_{k-1}\| \\
&\leq \rho\|Y_k - Y_{k-1}\|,
\end{aligned} \tag{2.10}$$

where

$$\rho = 3(1+\eta)^2\|A\|^2 = 3(1+\eta)^2\frac{\eta}{(1+\eta)^3} = \frac{3\eta}{1+\eta} < 1. \tag{2.11}$$

Using (2.10) repeatedly, we have

$$\|Y_{k+1} - Y_k\| \le \rho^k,$$

which implies that the matrix sequence $\{Y_k\}$ is convergent. Let $\widetilde{Y} = \lim_{k\to\infty} Y_k$. Then $\widetilde{Y}$ is a solution to the matrix Equation (2.6) and satisfies $\|\widetilde{Y}\| \le 1+\eta$.

Notice that here it cannot be asserted that $\widetilde{Y}$ is Hermitian. In order to prove that $\widetilde{Y}$ is Hermitian, we consider the following matrix equation

$$Y = I + \frac{1}{2}(A^*Y^2AY + YA^*Y^2A). \tag{2.12}$$

It is not hard to verify that for any solution Y of the matrix Equation (2.6), no matter whether it is Hermitian or not, Y must satisfy the matrix Equation (2.12). In fact, if Y is a solution of the matrix Equation (2.6), then it follows that

$$(I - A^*Y^2A)Y = I,$$

which means that $I - A^*Y^2A$ is the inverse of Y, and so it is implied that

$$Y(I - A^*Y^2A) = I,$$

i.e.,

$$Y = YA^*Y^2A + I.$$

This, together with Y satisfying (2.6), implies that Y is a solution of the matrix Equation (2.12).

Similarly, define matrix sequence $\{Z_k\}$ by

$$Z_{k+1} = I + \frac{1}{2}(A^*Z_k^2AZ_k + Z_kA^*Z_k^2A),$$

where $Z_0 = 0$. Clearly, such Z_k are all Hermitian. Moreover, following the same lines as the proof of (2.10) it can be proved that

$$\|Z_k\| \le 1+\eta \quad \textit{and} \quad \|Z_{k+1} - Z_k\| \le \rho^k$$

for all $k = 1, 2, \ldots$, where η and ρ defined by (2.5) and (2.11), respectively. Thus there must be a Hermitian matrix Z with $\|Z\| \le 1+\eta$ such that $\lim_{k\to\infty} Z_k = Z$.

The next task is to prove that the matrix Equation (2.12) has a unique solution Y with $\|Y\| \leq 2-\eta$. Let Y be such a solution. Then it follows that

$$\begin{aligned}
\|Z_{k+1}-Y\| &= \frac{1}{2}\|A^*Z_k^2AZ_k + Z_kA^*Z_k^2A - A^*Y^2AY - YA^*Y^2A\| \\
&\leq \frac{1}{2}\|A\|^2\Big(2\|Z_k\|^2 + 2\|Z_k\|\|Y\| + 2\|Y\|^2\Big)\|Z_k - Y\| \\
&\leq \|A\|^2\big((1+\eta)^2 + (1+\eta)(2-\eta) + (2-\eta)^2\big)\|Z_k - Y\| \\
&= \rho_1\|Z_k - Y\| \leq \rho_1^2\|Z_{k-1} - Y\| \leq \cdots \\
&\leq \rho_1^{k+1}\|Z_0 - Y\| = \rho_1^{k+1}\|Y\|,
\end{aligned}$$

where

$$\begin{aligned}
\rho_1 &= \|A\|^2\big((1+\eta)^2 + (1+\eta)(2-\eta) + (2-\eta)^2\big) \\
&= \|A\|^2(\eta^2 - \eta + 7) = 1 - \frac{(2\eta-1)^2}{(1+\eta)^3} < 1,
\end{aligned}$$

and hence, we have

$$Y = \lim_{k\to\infty} Z_k = Z.$$

In particular, it follows that

$$\widetilde{Y} = Z \quad and \quad \|Y\| > 2-\eta$$

for any other solution Y of the matrix Equation (2.6). Thus $\widetilde{Y}$ must be a Hermitian solution of the matrix Equation (2.6).

Notice that

$$\|A^*\widetilde{Y}^2A\widetilde{Y}\| \leq \|A\|^2\|\widetilde{Y}\|^3 \leq \|A\|^2(1+\eta)^3 = \eta < \frac{1}{2},$$

then we have

$$\widetilde{Y} = I + A^*\widetilde{Y}^2A\widetilde{Y} \geq (1-\frac{1}{2})I = \frac{1}{2}I,$$

which means that $\widetilde{Y}$ is a Hermitian positive definite matrix. Moreover, $\|\widetilde{Y}\| \leq 1+\eta < \frac{3}{2}$ implies that $\widetilde{Y} < \frac{3}{2}I$, and so $\widetilde{Y}^{-1} > \frac{2}{3}I$. Consequently, let $X_L = \widetilde{Y}^{-1}$. Then we complete the proof of the theorem. ■

Remark 2.1. Under the condition $\|A\| < \frac{2\sqrt{3}}{9}$, Hasanov and Ivanov [5] showed that the matrix Equation (1.1) has a h.p.d. solution X such that $\frac{2}{3}I < X \leq I$, and

Liu and Gao [6] proved such solution is unique. Here we used the different method to prove the same result and obtain some more useful properties of this solution.

In addition, it is worthwhile to point out that the condition $\|A\| < \frac{2\sqrt{3}}{9}$ is only a sufficient condition for which the matrix Equation (1.1) has the maximal solution X_L with $X_L > \frac{2}{3}I$. This can be seen from the following simple example.

Example 2.1. Let $A = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$, where $\frac{2\sqrt{3}}{9} \le \alpha < \frac{\sqrt{3}}{3}$. It is easy to verfy that $\|A\| = \alpha \ge \frac{2\sqrt{3}}{9}$, and $X_L = \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha^2 \end{bmatrix} \ge \frac{2}{3}I$ is the maximal solution of the matrix Equation (1.1) with this given data.

3. Perturbation Bound

Assume that the matrix Equation (1.1) has the maximal solution X_L with $X_L > \frac{2}{3}I$, and let the coefficient matrix A be slightly perturbed to $\widetilde{A} = A + \Delta A$. We next derive a perturbation bound for X_L.

Let $\widetilde{X}_L = X_L + \Delta X$ with $\Delta X \in \mathbf{H}^{n\times n}$ satisfies the perturbed matrix equation

$$\widetilde{X}_L + \widetilde{A}^* \widetilde{X}_L^{-2} \widetilde{A} = I. \tag{3.1}$$

Since X_L satisfies

$$X_L + A^* X_L^{-2} A = I, \tag{3.2}$$

subtracting (3.2) from (3.1) gives rise to

$$\Delta X - A^*(X_L^{-2} - \widetilde{X}_L^{-2})A + A^* \widetilde{X}_L^{-2} \Delta A + (\Delta A)^* \widetilde{X}_L^{-2} \widetilde{A} = 0. \tag{3.3}$$

Noting that

$$\begin{aligned} X_L^{-2} - \widetilde{X}_L^{-2} &= X_L^{-2} \Delta X \widetilde{X}_L^{-1} + X_L^{-1} \Delta X \widetilde{X}_L^{-2} \\ &= X_L^{-2} \Delta X \left(X_L^{-1} - X_L^{-1} \Delta X \widetilde{X}_L^{-1}\right) + X_L^{-1} \Delta X \left(X_L^{-1} - X_L^{-1} \Delta X \widetilde{X}_L^{-1}\right)^2 \\ &= X_L^{-2} \Delta X X_L^{-1} - X_L^{-2} \Delta X X_L^{-1} \Delta X \widetilde{X}_L^{-1} + X_L^{-1} \Delta X X_L^{-2} \\ &\quad + X_L^{-1} \Delta X \left[(X_L^{-1} \Delta X \widetilde{X}_L^{-1})^2 - X_L^{-2} \Delta X \widetilde{X}_L^{-1} - X_L^{-1} \Delta X \widetilde{X}_L^{-1} X_L^{-1}\right], \end{aligned}$$

Equation (3.3) can be written as

$$\mathbf{L}(\Delta X) = \mathbf{Q}(\Delta A) + \mathbf{F}(\Delta X), \tag{3.4}$$

where

$$\mathbf{L}(\Delta X) = \Delta X - (X_L^{-2}A)^*\Delta X(X_L^{-1}A) - (X_L^{-1}A)^*\Delta X(X_L^{-2}A), \tag{3.5}$$

$$\mathbf{Q}(\Delta A) = -A^*\widetilde{X}_L^{-2}\Delta A - (\Delta A)^*\widetilde{X}_L^{-2}A - (\Delta A)^*\widetilde{X}_L^{-2}\Delta A, \tag{3.6}$$

$$\begin{aligned}\mathbf{F}(\Delta X) =& -A^*X_L^{-2}\Delta X X_L^{-1}\Delta X\widetilde{X}_L^{-1}A + A^*X_L^{-1}\Delta X\big[(X_L^{-1}\Delta X\widetilde{X}_L^{-1})^2 \\ & -X_L^{-2}\Delta X\widetilde{X}_L^{-1} - X_L^{-1}\Delta X\widetilde{X}_L^{-1}X_L^{-1}\big]A.\end{aligned} \tag{3.7}$$

Next we show that the operator $\mathbf{L}$ is invertible.

Lemma 3.1. *The linear operator* $\mathbf{L}$ *defined by (3.5) is invertible, and* $\|\mathbf{L}^{-1}\| \le \frac{1}{3-2\zeta}$, *where* $\zeta = \|X_L^{-1}\|$.

Proof. Let $X_L = V\Gamma V^*$ be the spectral decomposition of X_L. By Lemma 2.1, there exist unitary matrix U and diagonal matrix $\Sigma \ge 0$ such that $\Gamma + \Sigma^2 = I$ and $A = V\Gamma U\Sigma V^*$. Thus, it follows from $\zeta = \|X_L^{-1}\|$ that $\|\Gamma^{-1}\| = \zeta$ and $\|\Sigma\|^2 = 1 - \frac{1}{\zeta}$, and hence we have

$$\begin{aligned}\|\mathbf{L}(W)\| &\ge \|W\| - 2\|X_L^{-1}A\|\|X_L^{-2}A\|\|W\| \\ &= \|W\|\big(1 - 2\|\Sigma\|\|\Gamma^{-1}U\Sigma\|\big) \\ &\ge \|W\|\big(1 - 2\|\Sigma\|^2\|\Gamma^{-1}\|\big) \\ &= (3-2\zeta)\|W\|\end{aligned} \tag{3.8}$$

for any $W \in \mathbf{H}^{n\times n}$. Since $\zeta = \|X_L^{-1}\| < \frac{3}{2}$, (3.8) implies that $\mathbf{L}$ is invertible and $\|\mathbf{L}^{-1}\| \le \frac{1}{3-2\zeta}$. ■

Applying Lemma 3.1 we can rewrite (3.4) as

$$\Delta X = \mathbf{L}^{-1}\big(\mathbf{Q}(\Delta A)\big) + \mathbf{L}^{-1}\big(\mathbf{F}(\Delta X)\big). \tag{3.9}$$

Now let

$$\epsilon = \|\Delta A\|, \quad \zeta = \|X_L^{-1}\|, \quad l = \|\mathbf{L}^{-1}\|^{-1}, \quad t = \frac{\sqrt{\zeta-1}}{\zeta\sqrt{\zeta}}, \tag{3.10}$$

and define

$$\alpha = \frac{(3-2\zeta)[4l\zeta^2 - (3-2\zeta)(4\zeta^2-\zeta-3)]}{27\zeta^3}, \qquad \gamma = \frac{l}{3\zeta(l+2\zeta-2)}. \tag{3.11}$$

Then, by Lemma 3.1 and the assumption for X_L, we have

$$1 \le \zeta = \|X_L^{-1}\| < \frac{2}{3} \quad \text{and} \quad l = \|\mathbf{L}^{-1}\|^{-1} \ge 3 - 2\zeta, \tag{3.12}$$

and so

$$\alpha \ge \frac{(3-2\zeta)[4(3-2\zeta)\zeta^2 - (3-2\zeta)(4\zeta^2 - \zeta - 3)]}{27\zeta^3} = \frac{(3-2\zeta)^2(\zeta+3)}{27\zeta^3} > 0.$$

The main result of this section is as follows:

Theorem 3.1. *If* $0 < \epsilon < \sqrt{\alpha + t^2} - t$, *then the maximal solution* $\widetilde{X}_L$, *with* $\widetilde{X}_L > \frac{2}{3}I$, *of the perturbed matrix Equation* (3.1) *exists, and we have*

$$\|\widetilde{X}_L - X_L\| \le \delta_*, \tag{3.13}$$

where δ_* *is the minimal positive solution of the equation*

$$\zeta^2(l + 2\zeta - 2)\delta^3 - \zeta(2l + 3\zeta - 3)\delta^2 + l\delta - \zeta^2(2t\epsilon + \epsilon^2) = 0. \tag{3.14}$$

Proof. Let

$$f(\Delta X) = \mathbf{L}^{-1}\big(\mathbf{Q}(\Delta A)\big) + \mathbf{L}^{-1}\big(\mathbf{F}(\Delta X)\big).$$

Obviously, $f(\Delta X)$ can be regarded as a continuous mapping from $\mathbf{H}^{n\times n}$ to $\mathbf{H}^{n\times n}$.

Define

$$\beta = \frac{l^2(4l + 9\zeta - 9)}{27\zeta^3(l + 2\zeta - 2)^2}, \tag{3.15}$$

and let

$$g(x) = \zeta^2(l + 2\zeta - 2)x^3 - \zeta(2l + 3\zeta - 3)x^2 + lx - \zeta^2(2t\epsilon + \epsilon^2). \tag{3.16}$$

Then we have

$$\begin{aligned} g'(x) &= 3\zeta^2(l + 2\zeta - 2)x^2 - 2\zeta(2l + 3\zeta - 3)x + l \\ &= \big(3\zeta(l + 2\zeta - 2)x - l\big)(\zeta x - 1) \\ &= 3\zeta^2(l + 2\zeta - 2)(x - \gamma)(x - \frac{1}{\zeta}), \end{aligned}$$

where γ defined as in (3.11). Since $\gamma < \frac{1}{\zeta}$, then it follows that $g(x)$ is strictly monotonically increasing on $(-\infty,\, \gamma] \cup [\frac{1}{\zeta},\, \infty)$, and strictly monotonically decreasing on $[\gamma,\, \frac{1}{\zeta}]$. Hence,

$$g(\gamma) = \frac{l^2(4l + 9\zeta - 9)}{27\zeta(l + 2\zeta - 2)^2} - \zeta^2(\epsilon^2 + 2t\epsilon) = -\zeta^2(\epsilon^2 + 2t\epsilon - \beta) \tag{3.17}$$

is a maximal value of $g(x)$, and $g(\frac{1}{\zeta}) = \frac{1-\zeta}{\zeta} - \zeta^2(\epsilon^2 + 2t\epsilon) < 0$ is a minimal value of $g(x)$, where β defined as in (3.15). Moreover, it is easy to verify that $\gamma \geq \dfrac{1}{\zeta} - \dfrac{2}{3}$, so $g(\gamma) \geq g\big(\dfrac{1}{\zeta} - \dfrac{2}{3}\big)$. On the other hand, if $0 < \epsilon < \sqrt{t^2 + \alpha} - t$, we have

$$
\begin{aligned}
g\big(\frac{1}{\zeta} - \frac{2}{3}\big) &= \zeta^2(l + 2\zeta - 2)\Big(\frac{1}{\zeta} - \frac{2}{3}\Big)^3 - \zeta(2l + 3\zeta - 3)\Big(\frac{1}{\zeta} - \frac{2}{3}\Big)^2 \\
&\quad + l\Big(\frac{1}{\zeta} - \frac{2}{3}\Big) - \zeta^2(\epsilon^2 + 2t\epsilon) \\
&= \frac{3-2\zeta}{3\zeta}\Big[(l+2\zeta-2)\Big(1-\frac{4}{3}\zeta+\frac{4}{9}\zeta^2\Big)-(2l+3\zeta-3)\Big(1-\frac{2}{3}\zeta\Big)+l\Big] \\
&\quad -\zeta^2(\epsilon^2 + 2t\epsilon) \\
&= \frac{3-2\zeta}{27\zeta}(4l\zeta^2 + 9 + 8\zeta^3 - 3\zeta - 14\zeta^2) - \zeta^2(\epsilon^2 + 2t\epsilon) \\
&= \frac{3-2\zeta}{27\zeta}\big[4l\zeta^2 - (3 - 2\zeta)(4\zeta^2 - \zeta - 3)\big] - \zeta^2(\epsilon^2 + 2t\epsilon) \\
&= -\zeta^2(\epsilon^2 + 2t\epsilon - \alpha) > 0.
\end{aligned}
\tag{3.18}
$$

This, together with $g(0) < 0$ and $g(\frac{1}{\zeta}) < 0$, shows that equation (3.14) has three positive real roots x_1, x_2, x_3 such that $0 < x_1 < \frac{1}{\zeta} - \frac{2}{3} \leq \gamma < x_2 < \frac{1}{\zeta} < x_3$. We use δ_* to denote the smallest one, that is, $\delta_* = x_1$.

Now define

$$
\mathcal{S}_{\delta_*} = \{\Delta X \in \mathbf{H}^{n\times n} \ : \ \|\Delta X\| \leq \delta_*\}. \tag{3.19}
$$

Then for any $\Delta X \in \mathcal{S}_{\delta_*}$ we have

$$
\|X_L^{-1}\Delta X\| \leq \|X_L^{-1}\|\|\Delta X\| \leq \zeta\delta_* < 1.
$$

Hence the matrix $I + X_L^{-1}\Delta X$ is nonsingular and

$$
\begin{aligned}
\|(X_L + \Delta X)^{-1}\| &= \|(I + X_L^{-1}\Delta X)^{-1}X_L^{-1}\| \\
&\leq \tfrac{\|X_L^{-1}\|}{1-\|X_L^{-1}\Delta X\|} \leq \tfrac{\zeta}{1-\zeta\|\Delta X\|}.
\end{aligned}
\tag{3.20}
$$

Noting that the proof of Lemma 3.1 implies that

$$
\|X_L^{-1}A\| = \|\Sigma\| = (1 - \frac{1}{\zeta})^{\frac{1}{2}}, \qquad \|X_L^{-2}A\| = \|\Gamma^{-1}U\Sigma\| \leq \zeta(1 - \frac{1}{\zeta})^{\frac{1}{2}},
$$

we have

$$
\|(X_L + \Delta X)^{-1}A\| = \|(I + X_L^{-1}\Delta X)^{-1}X_L^{-1}A\| \leq \frac{(1 - \frac{1}{\zeta})^{\frac{1}{2}}}{1 - \zeta\|\Delta X\|}, \tag{3.21}
$$

(3.22) $$\|(X_L+\Delta X)^{-2}A\| = \|(I+X_L^{-1}\Delta X)^{-1}X_L^{-1}(I+X_L^{-1}\Delta X)^{-1}X_L^{-1}A\| \le \frac{\zeta(1-\frac{1}{\zeta})^{\frac{1}{2}}}{(1-\zeta\|\Delta X\|)^2}.$$

Combining (3.6), (3.7), (3.10), (3.11), (3.20), (3.21), and (3.22), we obtain

$$\begin{aligned}\|f(\Delta X)\| \le& \frac{1}{l}\|\mathbf{Q}(\Delta A)\| + \frac{1}{l}\|\mathbf{F}(\Delta X)\| \\ \le& \frac{1}{l}\Bigg\{\frac{2\epsilon\zeta(1-\frac{1}{\zeta})^{\frac{1}{2}}}{(1-\zeta\|\Delta X\|)^2} + \frac{\epsilon^2\zeta^2}{(1-\zeta\|\Delta X\|)^2} + \frac{\zeta^2\|\Delta X\|^2(1-\frac{1}{\zeta})}{1-\zeta\|\Delta X\|} \\ &+ \|\Delta X\|(1-\frac{1}{\zeta})^{\frac{1}{2}}\left(\frac{\zeta^3\|\Delta X\|^2(1-\frac{1}{\zeta})^{\frac{1}{2}}}{(1-\zeta\|\Delta X\|)^2} + 2\frac{\zeta^2\|\Delta X\|(1-\frac{1}{\zeta})^{\frac{1}{2}}}{1-\zeta\|\Delta X\|}\right)\Bigg\} \\ =& \frac{\zeta^2(\epsilon^2+2t\epsilon)+3(\zeta^2-\zeta)\|\Delta X\|^2-2(\zeta^3-\zeta^2)\|\Delta X\|^3}{l(1-\zeta\|\Delta X\|)^2} \\ \le& \frac{\zeta^2(\epsilon^2+2t\epsilon)+3(\zeta^2-\zeta)\delta_*^2-2(\zeta^3-\zeta^2)\delta_*^3}{l(1-\zeta\delta_*)^2} = \delta_*\end{aligned}$$

for any $\Delta X \in \mathcal{S}_{\delta_*}$, in which the last equality is due to the fact that δ_* is a solution to the Equation (3.14). Thus we have proved that $f(\mathcal{S}_{\delta_*}) \subset \mathcal{S}_{\delta_*}$. By the Schauder fixed-point theorem, there exists a $\Delta X_* \in \mathcal{S}_{\delta_*}$ such that $f(\Delta X_*) = \Delta X_*$, i.e., there exists a solution ΔX_* to the perturbed Equation (3.4) such that

(3.23) $$\|\Delta X_*\| \le \delta_*.$$

Let $\widetilde{X}_L = X_L + \Delta X_*$. Then $\widetilde{X}_L$ is a Hermitian solution of the perturbed matrix Equation (3.1). Noting that $\|X_L^{-1}\|\|\Delta X_*\| < 1$ and that X_L is a Hermitian positive definite matrix, we know that $\widetilde{X}_L$ is also Hermitian positive definite. Since $\delta_* < \dfrac{1}{\zeta} - \dfrac{2}{3}$, we have

$$\begin{aligned}\|\widetilde{X}_L^{-1}\| &= \|(X_L+\Delta X_*)^{-1}\| = \|(I+X_L^{-1}\Delta X_*)^{-1}X_L^{-1}\| \\ &\le \frac{\|X_L^{-1}\|}{1-\|X_L^{-1}\|\|\Delta X_*\|} \le \frac{\zeta}{1-\zeta\delta_*} < \frac{3}{2},\end{aligned}$$

which implies that $\widetilde{X}_L^{-1} < \frac{3}{2}I$, and so, $\widetilde{X}_L > \frac{2}{3}I$. Thus, inequality (3.23) is just inequality (3.13). The proof is completed. ■

Remark 3.1. Consider $F(x,\epsilon) \equiv g(x)$, where $g(x)$ is defined by (3.16). It is easy to verify that $F(0,0) = 0$ and $\frac{\partial F}{\partial x}(0,0) \ne 0$. By the Implicit Function Theorem, there exists an analytic function $\delta_*(\epsilon) : \mathcal{N}_\epsilon \longrightarrow \mathbf{R}$ such that $F(\delta_*(\epsilon), \epsilon) \equiv 0$ for any $\epsilon \in \mathcal{N}_\epsilon$, where $\mathcal{N}_\epsilon$ is a neighborhood of the origin. It is easy to see that the $\delta_* = \delta_*(\epsilon)$ is just the minimal positive solution of Equation (3.14) for sufficient small positive number ϵ. Hence, we can get the second order absolute perturbation bound for the maximal solution X_L as follows:

$$\|\widetilde{X}_L - X_L\| \leq \delta_*'(0)\|\Delta A\| + \frac{1}{2}\delta_*''(0)\|\Delta A\|^2 + O(\|\Delta A\|^3), \quad \Delta A \longrightarrow 0, \tag{3.24}$$

where

$$\delta_*'(0) = \frac{2t\zeta^2}{l}, \quad \delta_*''(0) = \frac{2\zeta^2 l^2 + 8\zeta^2(\zeta-1)(2l+3\zeta-3)}{l^3}.$$

Combining this with (3.9) gives rise to

$$\Delta X = -\mathbf{L}^{-1}\big(A^* X_L^{-2}\Delta A + (\Delta A)^* X_L^{-2} A\big) + O(\|\Delta A\|^2), \quad \Delta A \longrightarrow 0. \tag{3.25}$$

4. Condition Numbers

We now apply the theory of condition developed by Rice [8] to study condition numbers of the maximal solution X_L, with $X_L > \frac{2}{3}I$, to the matrix Equation (1.1).

Suppose that the coefficient matrix A is slightly perturbed to $\widetilde{A} \in \mathbf{C}^{n\times n}$, and let $\Delta A = \widetilde{A} - A$. From Theorem 3.1 and Remark 3.1 we see that if $\|\Delta A\|_F$ is sufficiently small, then the maximal solution $\widetilde{X}_L$ to the perturbed matrix Equation (3.1) exists, and

$$\Delta X \equiv \widetilde{X}_L - X_L = -\mathbf{L}^{-1}\big(A^* X_L^{-2}\Delta A + (\Delta A)^* X_L^{-2} A\big) + O(\|\Delta A\|_F^2), \tag{4.1}$$

as $\Delta A \longrightarrow 0$.

By the theory of condition developed by Rice [8] we define the condition number of the maximal solution X_L by

$$c(X_L) = \lim_{\delta \to 0} \sup_{\|\frac{\Delta A}{\alpha}\|_F \leq \delta} \frac{\|\Delta X\|_F}{\xi\delta}, \tag{4.2}$$

where ξ and α are positive parameters. Taking $\xi = \alpha = 1$ in (4.2) gives the absolute condition number $c_{abs}(X_L)$, and taking $\xi = \|X_L\|_F$, $\alpha = \|A\|_F$ in (4.2) gives the relative condition number $c_{rel}(X_L)$.

Substituting (4.1) into (4.2) we get

$$\begin{aligned} c(X_L) &= \frac{1}{\xi} \max_{\substack{\frac{\Delta A}{\alpha} \neq 0 \\ \Delta A \in \mathbf{C}^{n\times n}}} \frac{\|\mathbf{L}^{-1}(C^*\Delta A + (\Delta A)^* C)\|_F}{\|\frac{\Delta A}{\alpha}\|_F} \\ &= \frac{1}{\xi} \max_{\substack{E \neq 0 \\ E \in \mathbf{C}^{n\times n}}} \frac{\|\alpha\mathbf{L}^{-1}(C^* E + E^* C)\|_F}{\|E\|_F}, \end{aligned} \tag{4.3}$$

where $C = X_L^{-2}A$.

Let L be the matrix representation of the linear operator $\mathbf{L}$, then it follows from (3.5) that

$$L = I \otimes I - D^T \otimes C^* - C^T \otimes D^*, \tag{4.4}$$

where $D = X_L^{-1}A.$, and let

$$\begin{aligned} w = \mathrm{vec}(E) &= u + iv, \quad g = (u^T, v^T)^T, \quad u, v \in \mathbf{R}^{n^2}, \\ L^{-1}(I \otimes C^*) &= L^{-1}\big(I \otimes (X_L^{-2}A)^*\big) = U_1 + i\Omega_1, \\ L^{-1}(C^T \otimes I)\Pi &= L^{-1}\big((X_L^{-2}A)^T \otimes I\big)\Pi = U_2 + i\Omega_2, \end{aligned} \tag{4.5}$$

where U_1, U_2, Ω_1, $\Omega_2 \in \mathbf{R}^{n^2 \times n^2}$, and Π is the vec-permutation matrix [3], i.e.,

$$\mathrm{vec}(E^T) = \Pi \mathrm{vec}(E).$$

Moreover, let

$$U_c = \begin{bmatrix} U_1 + U_2 & \Omega_2 - \Omega_1 \\ \Omega_1 + \Omega_2 & U_1 - U_2 \end{bmatrix}, \tag{4.6}$$

then from (4.3) we get

$$\begin{aligned} c(X_L) &= \frac{\alpha}{\xi} \max_{\substack{E \neq 0 \\ E \in \mathbf{C}^{n \times n}}} \frac{\|L^{-1}(I \otimes C^*)\mathrm{vec}(E) + L^{-1}(C^T \otimes I)\mathrm{vec}(E^*)\|}{\|\mathrm{vec}(E)\|} \\ &= \frac{\alpha}{\xi} \max_{w \neq 0} \frac{\|(U_1 + i\Omega_1)w + (U_2 + i\Omega_2)\bar{w}\|}{\|w\|} \\ &= \frac{\alpha}{\xi} \max_{g \neq 0} \frac{\|U_c g\|}{\|g\|} = \frac{\alpha}{\xi}\|U_c\|. \end{aligned} \tag{4.7}$$

Overall, we have the following result.

Theorem 4.1. *The condition number* $c(X_L)$ *defined by* (4.2) *has the explicit expression*

$$c(X_L) = \frac{\alpha}{\xi}\|U_c\|, \tag{4.8}$$

where the matrix U_c *is defined by* (4.6).

Remark 4.1. From (4.8) we have the relative condition number

$$c_{rel}(X_L) = \frac{\|A\|_F \|U_c\|}{\|X_L\|_F}.$$

5. Numerical Examples

To illustrate the results of the previous sections, in this section some simple examples are given, which were carried out using MATLAB 6.5 on a PC Pentium IV/1.7G computer, with machine epsilon $\epsilon \approx 2.2 \times 10^{-16}$.

Example 5.1. Consider the matrix equation

$$X + A_k^* X^{-2} A_k = I,$$

with $A_k = \frac{\delta_k}{\|A\|_2} A$, where

$$\delta_k = \frac{2\sqrt{3}}{9} - 10^{-k}, \quad A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

We now consider perturbation bounds for the maximal solution $X_L^{(k)}$ when the coefficient matrix A_k is perturbed to $A_{kj} = A_k + 10^{-j} A_0$, where A_0 is a random matrix generated by MATLAB function **rand** with $\|A_0\| = 1$. By Theorem 3.1, we can compute the perturbation bounds $\delta_*^{(kj)}$ with the perturbed matrices A_{kj}:

$$r^{(kj)} \equiv \|X_L^{(k)} - X_L^{(kj)}\| \leq \delta_*^{(kj)},$$

where $X_L^{(k)}$ and $X_L^{(kj)}$ are the maximal solutions with the coefficient matrices A_k and A_{kj}, respectively. Some results are listed in Table 5.1.

From the results listed in Table 5.1 we see that the perturbation bound of X_L decreases as the error $\|\widetilde{A} - A\|$ decreases, and moreover, we can see that our estimate for the perturbation bound is quite good in such cases.

Table 5.1

j	5	6	7	8	9
$r^{(1j)}$	4.847×10^{-6}	4.847×10^{-7}	4.847×10^{-8}	4.847×10^{-9}	4.842×10^{-10}
$\delta_*^{(1j)}$	3.581×10^{-4}	3.557×10^{-5}	3.555×10^{-6}	3.554×10^{-7}	3.554×10^{-8}
$r^{(2j)}$	1.601×10^{-5}	1.601×10^{-6}	1.601×10^{-7}	1.601×10^{-8}	1.596×10^{-9}
$\delta_*^{(2j)}$	1.205×10^{-4}	1.203×10^{-5}	1.203×10^{-6}	1.203×10^{-7}	1.203×10^{-8}
$r^{(3j)}$	4.801×10^{-5}	4.797×10^{-6}	4.797×10^{-7}	4.795×10^{-8}	4.781×10^{-9}
$\delta_*^{(3j)}$	3.769×10^{-4}	3.709×10^{-5}	3.703×10^{-6}	3.703×10^{-7}	3.702×10^{-8}

Example 5.2. Consider the matrix Equation (1.1) with the coefficient matrix $A = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$, where $0 < \alpha < \frac{\sqrt{3}}{3}$. In this case the matrix Equation (1.1) has the maximal solution

$$X_L = \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha^2 \end{bmatrix}.$$

Take $\alpha = \frac{2\sqrt{3}}{9} - 10^{-k}$, and suppose that the perturbation in the coefficient matrix A is

$$\Delta A = 10^{-10} \times \begin{bmatrix} 0.9501 & 0.6068 \\ 0.2311 & 0.4860 \end{bmatrix}.$$

Some numerical results on the relative perturbation bounds $\delta_*/\|X_L\|_F$, $\widetilde{\delta}$ and $c_{rel}(X_L)$ are shown in Table 5.2, where δ_* is as in (3.13) with the unitary invariant norm $\|\cdot\|_F$, $\widetilde{\delta}$ is the relative perturbation bound given by Proposition 4.1 in [7], and $c_{rel}(X_L)$ is given in Remark 4.1.

Table 5.2

k	2	4	6	8	10
$\delta_*/\|X_L\|_F$	0.841×10^{-8}	0.788×10^{-8}	0.787×10^{-8}	0.787×10^{-8}	0.787×10^{-8}
$\widetilde{\delta}$	0.266×10^{-8}	0.273×10^{-6}	0.273×10^{-4}	0.274×10^{-2}	0.392
$c_{rel}(X_L)$	0.2132	0.2254	0.2256	0.2256	0.2256

The results listed in Table 5.2 show that the relative perturbation bound $\delta_*/\|X_L\|_F$ is fairly sharp, while the bound $\widetilde{\delta}$ given by [7] is conservative.

On the other hand, take $\alpha = 0.57 < \frac{\sqrt{3}}{3}$, and suppose that the perturbation in the coefficient matrix is

$$\Delta A = 10^{-k} \times \begin{bmatrix} 0.9501 & 0.6068 \\ 0.2311 & 0.4860 \end{bmatrix}.$$

In this case the relative condition number is $c_{rel}(X_L) = 0.5386$, which is computed by the formula given as in Remark 4.1. This shows that the maximal solution X_L is well-conditioned. Since the condition (b) of Proposition 4.1 in [7] is violated, $\widetilde{\delta}$ becomes negative, and so we can not use it as a perturbation bound. However, as shown in Table 5.3, in such a case $\delta_*/\|X_L\|_F$ can still give quite sharp perturbation bounds.

Table 5.3

k	6	7	8	9	10
$\delta_*/\|X_L\|_F$	3.404×10^{-5}	3.398×10^{-6}	3.398×10^{-7}	3.398×10^{-8}	3.398×10^{-9}

REFERENCES

1. S. M. El-Sayed, Two iteration processes for computing positive definite solutions of the equation $X - A^*X^{-n}A = Q$, *Computers and Mathematics with Applications*, **41** (2001), 579-588.
2. S. M. El-Sayed and A. C. M. Ran, On an iteration method for solving a class of nonlinear matrix equations, *SIAM J. Matrix Anal. Appl.*, **23** (2001), 632-645.
3. A. Graham, *Kronecker products and matrix calculus: with applications*, Ellis Horwood Limited, Chichester, 1981.
4. V. I. Hasanov and I. G. Ivanov, Positive definite solutions of the equation $X+A^*X^{-n}A = I$, *Lecture Notes in Computer Science, Numerical Analysis and its Applications* 2000, 377-384 (2001).
5. I. G. Ivanov, V. I. Hasanov and B.V. Minchev, On matrix equations $X \pm A^*X^{-2}A = I$, *Linear Algebra Appl.*, **326** (2001), 27-44.
6. X.-G. Liu and H. Gao, On the positive definite solutions of the matrix equations $X^s \pm A^TX^{-t}A = I_n$, *Linear Algebra Appl.*, **368** (2003), 83-97.
7. A. C. M. Ran and M. C. B. Reurings, On the nonlinear matrix equation $X + A^*F(X)A = Q$: Solutions and perturbation theory, *Linear Algebra Appl.*, **346** (2002), 15–26.
8. J. R. Rice, A theory of condition, *J. SIAM Numer. Anal.*, **3** (1966), 287-310.
9. J.-G. Sun. and S. F. Xu, Perturbation analysis of the maximal solution of the matrix equation $X + A^*X^{-1}A = P$ II., *Linear Algebra Appl.*, **362** (2003), 211-228.
10. S. F. Xu, On the maximal solution of the matrix equation $X + A^TX^{-1}A = I$, *Acta Scientiarum Naturalium Universitatis Pekinensis*, **36** (2000), 29-38.
11. Y. H. Zhang, On Hermitian positive definite solutions of matrix equation $X+A^*X^{-2}A = I$, *Linear Algebra Appl.*, **372** (2003), 295-304.

Shufang Xu and Mingsong Cheng
School of Mathematical Sciences,
Peking University,
Beijing 100871,
People's Republic of China
E-mail: xsf@math.pku.edu.cn.
E-mail: mscheng@dlut.edu.cn.