Research Article

Sensitivity Analysis of the Matrix Equation from Interpolation Problems

Jing Li[1] and Yuhai Zhang[2]

[1] *School of Mathematics and Statistics, Shandong University, Weihai, Weihai 264209, China*
[2] *School of Mathematics, Shandong University, Jinan 250100, China*

Correspondence should be addressed to Jing Li; xlijing@sdu.edu.cn

This paper studies the sensitivity analysis of a nonlinear matrix equation connected to interpolation problems. The backward error estimates of an approximate solution to the equation are derived. A residual bound of an approximate solution to the equation is obtained. A perturbation bound for the unique solution to the equation is evaluated. This perturbation bound is independent of the exact solution of this equation. The theoretical results are illustrated by numerical examples.

1. Introduction

In this paper we consider the Hermitian positive definite solution of the nonlinear matrix equation:

$$X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q, \qquad (1)$$

where $A_1, A_2, \ldots, A_m$ are $n \times n$ complex matrices, m is a positive integer, and Q is a positive definite matrix. Here, A_i^* denotes the conjugate transpose of the matrix A_i.

This type of nonlinear matrix equations arises in many practical problems. The equation $X - A^* X^{-1} A = Q$ which is a representative of (1) for $m = 1$ comes from ladder networks, dynamic programming, control theory, stochastic filtering, statistics, and so forth [1–6]. When $m > 1$, (1) comes from the nonlinear matrix equation:

$$X = Q + A^* \left(\widehat{X} - C \right)^{-1} A, \qquad (2)$$

where Q is an $n \times n$ positive definite matrix, C is an $mn \times mn$ positive semidefinite matrix, A is an arbitrary $mn \times n$ matrix, and $\widehat{X}$ is the $m \times m$ block diagonal matrix with, on each diagonal entry, the $n \times n$ matrix X. In [7], (2) is recognized as playing an important role in modelling certain optimal interpolation problems. Let $C = 0$ and

$A = (A_1^T, A_2^T, \ldots, A_m^T)^T$, where A_i, $i = 1, 2, \ldots, m$, are $n \times n$ matrices. Then $X = Q + A^* (\widehat{X} - C)^{-1} A$ can be rewritten as $X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$. When we solve the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$, we often do not know A_i and Q exactly but have only approximations $\widetilde{A}_i$ and $\widetilde{Q}$ available. Then we only solve the equation $\widetilde{X} - \sum_{i=1}^{m} \widetilde{A}_i^* \widetilde{X}^{-1} \widetilde{A}_i = \widetilde{Q}$ exactly which gives a different solution $\widetilde{X}$. We would like to know how the errors of $\widetilde{A}_i$ and $\widetilde{Q}$ influence the error in $\widetilde{X}$. Motivated by this, we consider in this paper the sensitivity analysis of (1).

For the equation $X - A^* X^{-1} A = Q$ and related equations $X - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$ $(0 < |\delta_i| < 1)$ and $X - \sum_{i=1}^{m} A_i^* X^r A_i = Q$ $(-1 \le r < 0$ or $0 < r < 1)$ there were many contributions in the literature to the solvability and numerical solutions [8–14]. However, these papers have not examined the sensitivity analysis about the above equations. Hasanov et al. [12, 15] obtained two perturbation estimates of the solutions to the equations $X \pm A^* X^{-1} A = Q$. Li and Zhang [13] derived two perturbation bounds of the unique solution to the equation $X - A^* X^{-1} A = Q$. They also obtained the explicit expression of the condition number for the unique positive definite solution. The perturbation analysis about the related equations $X + A^* X^{-1} A = P$, $X - A^* X^{-p} A = Q$, and $X^s \pm A^T X^{-t} A = I_n$ were mentioned in papers [16–19]. Yin and Fang [20] obtained an explicit expression of the condition

number for the unique positive definite solution of (1). They also gave two perturbation bounds for the unique positive definite solution, whereas, to our best knowledge, there have been no backward error estimates and any computable residual bound for (1) in the known literature. In this paper, we obtain the backward error estimates and a residual bound of the approximate solution to (1) as well as evaluate a new relative perturbation bound for (1). This bound does not need any knowledge of the exact solution of (1), which is important in many practical calculations.

As a continuation of the previous results, this paper gives some preliminary knowledge that will be needed to develop this work in Section 2. In Section 3, the backward error estimates of an approximate solution for the unique solution to (1) are discussed. In Section 4, we derive a residual bound of an approximate solution for the unique solution to (1). In Section 5, we give a new perturbation bound for the unique solution to (1), which is independent of the exact solution of (1). Finally, several numerical examples are presented in Section 6.

We denote by $\mathscr{C}^{n\times n}$ the set of $n \times n$ complex matrices, by $\mathscr{H}^{n\times n}$ the set of $n \times n$ Hermitian matrices, by I the identity matrix, by $\mathbf{i}$ the imaginary unit, by $\|\cdot\|$ the spectral norm, by $\|\cdot\|_F$ the Frobenius norm, and by $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$ the maximal and minimal eigenvalues of M, respectively. For $A = (a_1, \ldots, a_n) = (a_{ij}) \in \mathscr{C}^{n\times n}$ and a matrix B, $A \otimes B = (a_{ij}B)$ is a Kronecker product, and $\text{vec } A$ is a vector defined by $\text{vec } A = (a_1^T, \ldots, a_n^T)^T$. For $X, Y \in \mathscr{H}^{n\times n}$, we write $X \geq Y$ (resp., $X > Y$) if $X - Y$ is Hermitian positive semidefinite (resp., definite).

2. Preliminaries

Lemma 1 (see [8, Theorem 2.1]). *The matrix equation*

$$X - \sum_{i=1}^{m} A_i^* X^r A_i = Q, \quad -1 \leq r < 0 \tag{3}$$

always has a unique positive definite solution.

Lemma 2 (see [8, Theorem 2.3]). *Let X be the unique Hermitian positive definite solution of (3); then $X \in [\beta I, \alpha I]$, where the pair (β, α) is a solution of the system*

$$\beta = \lambda_{\min}(Q) + \sum_{i=1}^{m} \lambda_{\min}(A_i^* A_i) \alpha^r,$$

$$\alpha = \lambda_{\max}(Q) + \sum_{i=1}^{m} \lambda_{\max}(A_i^* A_i) \beta^r. \tag{4}$$

3. Backward Error

In this section, applying the technique developed in [18], we obtain some estimates for the backward error of the approximate solution of (1).

Let $\widetilde{X} \in \mathscr{H}^{n\times n}$ be an approximation to the unique solution X to (1), and let $\Delta A_i \in \mathscr{C}^{n\times n}$ ($i = 1, 2, \ldots, m$) and $\Delta Q \in \mathscr{H}^{n\times n}$ be the corresponding perturbations of the coefficient

matrices A_i ($i = 1, 2, \ldots, m$) and Q in (1). A backward error of the approximate solution $\widetilde{X}$ can be defined by

$$\eta(\widetilde{X}) = \min \left\{ \left\| \left(\frac{\Delta A_1}{\alpha_1}, \frac{\Delta A_2}{\alpha_2}, \ldots, \frac{\Delta A_m}{\alpha_m}, \frac{\Delta Q}{\rho} \right) \right\|_F \right.$$

$$: \widetilde{X} - \sum_{i=1}^{m} (A_i + \Delta A_i)^* \widetilde{X}^{-1} (A_i + \Delta A_i) \tag{5}$$

$$\left. = Q + \Delta Q \right\},$$

where $\alpha_1, \alpha_2, \ldots, \alpha_m$ and ρ are positive parameters. Taking $\alpha_i = \|A_i\|_F$, $i = 1, 2, \ldots, m$, and $\rho = \|Q\|_F$ in (5) gives the relative backward error $\eta_{\text{rel}}(\widetilde{X})$, and taking $\alpha_i = 1$, $i = 1, 2, \ldots, m$, and $\rho = 1$ in (5) gives the absolute backward error $\eta_{\text{abs}}(\widetilde{X})$.

Let

$$R = Q - \widetilde{X} + \sum_{i=1}^{m} A_i^* \widetilde{X}^{-1} A_i. \tag{6}$$

Note that

$$Q = \widetilde{X} - \sum_{i=1}^{m} (A_i + \Delta A_i)^* \widetilde{X}^{-1} (A_i + \Delta A_i) - \Delta Q. \tag{7}$$

It follows from (6) that

$$-\sum_{i=1}^{m} \left(\Delta A_i^* \widetilde{X}^{-1} A_i + A_i^* \widetilde{X}^{-1} \Delta A_i \right) - \Delta Q$$

$$= R + \sum_{i=1}^{m} \Delta A_i^* \widetilde{X}^{-1} \Delta A_i. \tag{8}$$

Let

$$\left(I \otimes \left(\widetilde{X}^{-1} A_i \right)^* \right) = U_{i1} + \mathbf{i}\Omega_{i1},$$

$$\left(\left(\widetilde{X}^{-1} A_i \right)^T \otimes I \right) \Pi = U_{i2} + \mathbf{i}\Omega_{i2},$$

$$\text{vec } \Delta A_i = x_i + \mathbf{i}y_i, \qquad \text{vec } \Delta Q = q_1 + \mathbf{i}q_2,$$

$$\text{vec } R = r_1 + \mathbf{i}r_2,$$

$$\text{vec } \left(\Delta A_i^* \widetilde{X}^{-1} \Delta A_i \right) = a_i + \mathbf{i}b_i, \quad i = 1, 2, \ldots, m, \tag{9}$$

$$g = \left(\frac{x_1^T}{\alpha_1}, \frac{y_1^T}{\alpha_1}, \ldots, \frac{x_m^T}{\alpha_m}, \frac{y_m^T}{\alpha_m}, \frac{q_1^T}{\rho}, \frac{q_2^T}{\rho} \right)^T,$$

$$U_i = \begin{pmatrix} U_{i1} + U_{i2} & \Omega_{i2} - \Omega_{i1} \\ \Omega_{i1} + \Omega_{i2} & U_{i1} - U_{i2} \end{pmatrix},$$

$$T = \left[-\alpha_1 U_1, -\alpha_2 U_2, \ldots, -\alpha_m U_m, -\rho I_{2n^2} \right],$$

where Π is the vec-permutation. Then (8) can be written as

$$Tg = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} + \sum_{i=1}^{m} \begin{pmatrix} a_i \\ b_i \end{pmatrix}. \tag{10}$$

It follows from $\rho > 0$ that $2n^2 \times 2(m+1)n^2$ matrix T is full row rank. Hence, $TT^\dagger = I_{2n^2}$, which implies that every solution to the equation

$$g = T^\dagger \binom{r_1}{r_2} + T^\dagger \left(\sum_{i=1}^m \binom{a_i}{b_i} \right) \tag{11}$$

must be a solution to (10). Consequently, for any solution g to (11) we have

$$\eta\left(\widetilde{X}\right) \leq \|g\|. \tag{12}$$

Then we can state the estimates of the backward error as follows.

Theorem 3. *Let* $A_1, A_2, \ldots, A_m, Q, \widetilde{X} \in \mathscr{C}^{n\times n}$ *be given matrices and let* $\eta(\widetilde{X})$ *be the backward error defined by (5). If*

$$r < \frac{s}{4t \left(\sum_{i=1}^m \alpha_i^2 \right)}, \tag{13}$$

then one has that

$$U(r) \leq \eta\left(\widetilde{X}\right) \leq B(r), \tag{14}$$

where

$$r = \left\| T^\dagger \binom{r_1}{r_2} \right\|, \qquad s = \left\| T^\dagger \right\|^{-1}, \qquad t = \left\| \widetilde{X}^{-1} \right\|, \tag{15}$$

$$B(r) = \frac{2rs}{s + \sqrt{s^2 - 4rst \left(\sum_{i=1}^m \alpha_i^2 \right)}}, \tag{16}$$

$$U(r) = \frac{2r\sqrt{s^2 - 4rst \left(\sum_{i=1}^m \alpha_i^2 \right)}}{s + \sqrt{s^2 - 4rst \left(\sum_{i=1}^m \alpha_i^2 \right)}}.$$

Proof. Let

$$L(g) = T^\dagger \binom{r_1}{r_2} + T^\dagger \left(\sum_{i=1}^m \binom{a_i}{b_i} \right). \tag{17}$$

Obviously, $L : \mathscr{C}^{2(m+1)n^2\times 1} \rightarrow \mathscr{C}^{2(m+1)n^2\times 1}$ is continuous. Condition (13) ensures that the quadratic equation

$$x = r + \frac{t}{s} \left(\sum_{i=1}^m \alpha_i^2 \right) x^2 \tag{18}$$

in x has two positive real roots. The smaller one is

$$B(r) = \frac{2rs}{s + \sqrt{s^2 - 4rst \left(\sum_{i=1}^m \alpha_i^2 \right)}}. \tag{19}$$

Define $\Omega = \{g \in \mathscr{C}^{2(m+1)n^2\times 1} : \|g\| \leq B(r)\}$. Then for any $g \in \Omega$, we have

$$\begin{aligned}
\|L(g)\| &\leq r + \frac{1}{s} \sum_{i=1}^m \left\| \binom{a_i}{b_i} \right\| \\
&= r + \frac{1}{s} \sum_{i=1}^m \left\| \Delta A_i^* \widetilde{X}^{-1} \Delta A_i \right\|_F \leq r + \frac{t}{s} \sum_{i=1}^m \|\Delta A_i\|_F^2 \\
&\leq r + \frac{t}{s} \left(\sum_{i=1}^m \alpha_i^2 \right) \left\| \left(\frac{\Delta A_1}{\alpha_1}, \frac{\Delta A_2}{\alpha_2}, \ldots, \frac{\Delta A_m}{\alpha_m} \right) \right\|_F^2 \\
&\leq r + \frac{t}{s} \left(\sum_{i=1}^m \alpha_i^2 \right) \|g\|^2 \leq r + \frac{t}{s} \left(\sum_{i=1}^m \alpha_i^2 \right) B^2(r) = B(r).
\end{aligned} \tag{20}$$

The last equality is due to the fact that $B(r)$ is a solution to the quadratic equation (18). Thus we have proved that $L(\Omega) \subset \Omega$. By the Schauder fixed-point theorem, there exists a $g_* \in \Omega$ such that $L(g_*) = g_*$, which means that g_* is a solution to (11), and hence it follows from (12) that

$$\eta\left(\widetilde{X}\right) \leq \|g_*\| \leq B(r). \tag{21}$$

Next we derive a lower bound for $\eta(\widetilde{X})$. Suppose that $(\Delta A_{1\min}/\alpha_1, \ldots, \Delta A_{m\min}/\alpha_m, \Delta Q_{\min}/\rho)$ satisfies

$$\eta\left(\widetilde{X}\right) = \left\| \left(\frac{\Delta A_{1\min}}{\alpha_1}, \ldots, \frac{\Delta A_{m\min}}{\alpha_m}, \frac{\Delta Q_{\min}}{\rho} \right) \right\|_F. \tag{22}$$

Then we have

$$T g_{\min} = \binom{r_1}{r_2} + \sum_{i=1}^m \binom{a_{i\star}}{b_{i\star}}, \tag{23}$$

where

$$\begin{aligned}
&\mathrm{vec}\left(\Delta A_{i\min}^* \widetilde{X}^{-1} \Delta A_{i\min} \right) = a_{i\star} + \mathbf{i}b_{i\star}, \\
&\mathrm{vec}\left(\Delta A_{i\min} \right) = x_{i\star} + \mathbf{i}y_{i\star}, \\
&\mathrm{vec}\left(\Delta Q_{\min} \right) = q_{1\star} + \mathbf{i}q_{2\star}, \\
&g_{\min} = \left(\frac{x_{1\star}^T}{\alpha_1}, \frac{y_{1\star}^T}{\alpha_1}, \ldots, \frac{x_{m\star}^T}{\alpha_m}, \frac{y_{m\star}^T}{\alpha_m}, \frac{q_{1\star}^T}{\rho}, \frac{q_{2\star}^T}{\rho} \right)^T.
\end{aligned} \tag{24}$$

Let a singular value decomposition of T be $T = W(E, 0)Z^*$, where W and Z are unitary matrices and $E = \mathrm{diag}(e_1, e_2, \ldots, e_{2n^2})$ with $e_1 \geq \cdots \geq e_{2n^2} > 0$. Substituting this decomposition into (23), and letting

$$Z^* g_{\min} = \binom{v}{\star}, \qquad v \in \mathscr{C}^{2n^2\times 1}, \tag{25}$$

we get

$$v = E^{-1}W^* \binom{r_1}{r_2} + E^{-1}W^* \sum_{i=1}^m \binom{a_{i\star}}{b_{i\star}}. \tag{26}$$

It follows from (22) that

$$
\begin{aligned}
\eta\left(\widetilde{X}\right) &= \left\|g_{\min}\right\| = \left\|\binom{v}{\star}\right\| \geq \|v\| \\
&\geq \left\|E^{-1}W^*\binom{r_1}{r_2}\right\| - \left\|E^{-1}W^*\sum_{i=1}^{m}\binom{a_{i\star}}{b_{i\star}}\right\| \\
&\geq \left\|T^\dagger\binom{r_1}{r_2}\right\| - \left\|T^\dagger\right\| \cdot \sum_{i=1}^{m}\left\|\binom{a_{i\star}}{b_{i\star}}\right\| \\
&\geq r - \frac{1}{s}\sum_{i=1}^{m}\left\|\Delta A_{i\,\min}^*\widetilde{X}^{-1}\Delta A_{i\,\min}\right\|_F \\
&\geq r - \frac{t}{s}\sum_{i=1}^{m}\left\|\Delta A_{i\,\min}\right\|_F^2 \\
&\geq r - \frac{t}{s}\left(\sum_{i=1}^{m}\alpha_i^2\right)\left\|\left(\frac{\Delta A_{1\,\min}}{\alpha_1},\dots,\frac{\Delta A_{m\,\min}}{\alpha_m}\right)\right\|_F^2 \\
&\geq r - \frac{t}{s}\left(\sum_{i=1}^{m}\alpha_i^2\right)B^2(r).
\end{aligned}
\tag{27}
$$

Here we have used the fact that

$$
\begin{aligned}
&\left\|\left(\frac{\Delta A_{1\,\min}}{\alpha_1},\dots,\frac{\Delta A_{m\,\min}}{\alpha_m}\right)\right\|_F \\
&\leq \left\|\left(\frac{\Delta A_{1\,\min}}{\alpha_1},\dots,\frac{\Delta A_{m\,\min}}{\alpha_m},\frac{\Delta Q_{\min}}{\rho}\right)\right\|_F \\
&= \eta\left(\widetilde{X}\right) \leq B(r).
\end{aligned}
\tag{28}
$$

Let

$$
U(r) = r - \frac{t}{s}\left(\sum_{i=1}^{m}\alpha_i^2\right)B^2(r).
\tag{29}
$$

Since $B(r)$ is a solution to (18), we have that

$$
B(r) = r + \frac{t}{s}\left(\sum_{i=1}^{m}\alpha_i^2\right)B^2(r),
\tag{30}
$$

which implies that

$$
\begin{aligned}
U(r) &= r - \frac{t}{s}\left(\sum_{i=1}^{m}\alpha_i^2\right)B^2(r) = 2r - B(r) \\
&= \frac{2r\sqrt{s^2 - 4rst\left(\sum_{i=1}^{m}\alpha_i^2\right)}}{s + \sqrt{s^2 - 4rst\left(\sum_{i=1}^{m}\alpha_i^2\right)}} > 0.
\end{aligned}
\tag{31}
$$

Then $\eta(\widetilde{X}) \geq U(r)$.   □

4. Residual Bound

Residual bound reveals the stability of a numerical method. In this section, in order to derive the residual bound of an approximate solution for the unique solution to (1), we first introduce the following lemma.

Lemma 4. *For every positive definite matrix $X \in \mathscr{H}^{n\times n}$, if $X + \Delta X \geq (1/v)I > 0$, then*

$$
\begin{aligned}
&\left\|\sum_{i=1}^{m}A_i^*\left((X+\Delta X)^{-1} - X^{-1}\right)A_i\right\| \\
&\leq \left(\|\Delta X\| + v\|\Delta X\|^2\right)\sum_{i=1}^{m}\left\|X^{-1}A_i\right\|^2.
\end{aligned}
\tag{32}
$$

Proof. According to

$$
\begin{aligned}
&(X+\Delta X)^{-1} - X^{-1} \\
&= -X^{-1}\Delta X(X+\Delta X)^{-1} = -X^{-1}\Delta X X^{-1} \\
&\quad + X^{-1}\Delta X X^{-1}\Delta X(X+\Delta X)^{-1},
\end{aligned}
\tag{33}
$$

it follows that

$$
\begin{aligned}
&\left\|\sum_{i=1}^{m}A_i^*\left((X+\Delta X)^{-1} - X^{-1}\right)A_i\right\| \\
&\leq \sum_{i=1}^{m}\left(\left\|A_i^*X^{-1}\Delta X X^{-1}A_i\right\| \right.\\
&\quad\left. + \left\|A_i^*X^{-1}\Delta X X^{-1}\Delta X(X+\Delta X)^{-1}A_i\right\|\right) \\
&\leq \left(\|\Delta X\| + v\|\Delta X\|^2\right)\sum_{i=1}^{m}\left\|X^{-1}A_i\right\|^2. \quad\square
\end{aligned}
\tag{34}
$$

Theorem 5. *Let $\widetilde{X} > 0$ be an approximation to the solution X of (1). If the residual $R(\widetilde{X}) \equiv Q + \sum_{i=1}^{m}A_i^*\widetilde{X}^{-1}A_i - \widetilde{X}$ satisfies*

$$
\begin{aligned}
&\left\|R\left(\widetilde{X}\right)\right\| < \frac{(1-\Sigma)^2}{1+\Sigma+2\sqrt{\Sigma}}\,\lambda_{\min}\left(\widetilde{X}\right), \\
&\text{where } \Sigma \equiv \sum_{i=1}^{m}\left\|\widetilde{X}^{-1}A_i\right\|^2 < 1,
\end{aligned}
\tag{35}
$$

then

$$
\left\|\widetilde{X} - X\right\| \leq \theta\left\|R\left(\widetilde{X}\right)\right\|,
\tag{36}
$$

where

$$
\begin{aligned}
\theta = \;&\left(2\lambda_{\min}\left(\widetilde{X}\right)\right) \\
&\times\left((1-\Sigma)\lambda_{\min}\left(\widetilde{X}\right) + \left\|R\left(\widetilde{X}\right)\right\| \right.\\
&\quad + \left(\left((1-\Sigma)\lambda_{\min}\left(\widetilde{X}\right) + \left\|R\left(\widetilde{X}\right)\right\|\right)^2 \right.\\
&\quad\left.\left. - 4\lambda_{\min}\left(\widetilde{X}\right)\left\|R\left(\widetilde{X}\right)\right\|\right)^{1/2}\right)^{-1}.
\end{aligned}
\tag{37}
$$

Proof. Let

$$
\Psi = \left\{\Delta X \in \mathscr{H}^{n\times n}: \|\Delta X\| \leq \theta\left\|R\left(\widetilde{X}\right)\right\|\right\}.
\tag{38}
$$

Obviously, Ψ is a nonempty bounded convex closed set. Let

$$g(\Delta X) = \sum_{i=1}^{m} A_i^* \left[\left(\widetilde{X} + \Delta X \right)^{-1} - \widetilde{X}^{-1} \right] A_i + R\left(\widetilde{X} \right). \quad (39)$$

Evidently $g : \Psi \mapsto \mathcal{H}^{n \times n}$ is continuous.

Note that condition (35) ensures that the quadratical equation

$$x^2 - \left(\lambda_{\min}\left(\widetilde{X} \right)(1 - \Sigma) + \left\| R\left(\widetilde{X} \right) \right\| \right) x$$
$$+ \lambda_{\min}\left(\widetilde{X} \right) \left\| R\left(\widetilde{X} \right) \right\| = 0 \quad (40)$$

has two positive real roots, and the smaller one is given by

$$\mu_* = \left(2\lambda_{\min}\left(\widetilde{X} \right) \left\| R\left(\widetilde{X} \right) \right\| \right)$$
$$\times \left((1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right) + \left\| R\left(\widetilde{X} \right) \right\| \right.$$
$$+ \left(\left((1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right) + \left\| R\left(\widetilde{X} \right) \right\| \right)^2 \right. \quad (41)$$
$$\left. \left. -4\lambda_{\min}\left(\widetilde{X} \right) \left\| R\left(\widetilde{X} \right) \right\| \right)^{1/2} \right)^{-1}.$$

Next, we will prove that $g(\Psi) \subseteq \Psi$.

For every $\Delta X \in \Psi$, we have

$$\Delta X \geq -\theta \left\| R\left(\widetilde{X} \right) \right\| I. \quad (42)$$

Hence

$$\widetilde{X} + \Delta X \geq \widetilde{X} - \theta \left\| R\left(\widetilde{X} \right) \right\| I \geq \left(\lambda_{\min}\left(\widetilde{X} \right) - \theta \left\| R\left(\widetilde{X} \right) \right\| \right) I. \quad (43)$$

By (36), one sees that

$$\theta \left\| R\left(\widetilde{X} \right) \right\| \leq \frac{2\lambda_{\min}\left(\widetilde{X} \right) \left\| R\left(\widetilde{X} \right) \right\|}{(1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right) + \left\| R\left(\widetilde{X} \right) \right\|}$$
$$= \lambda_{\min}\left(\widetilde{X} \right) \quad (44)$$
$$\times \left(1 + \frac{\left\| R\left(\widetilde{X} \right) \right\| - (1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right)}{(1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right) + \left\| R\left(\widetilde{X} \right) \right\|} \right).$$

According to (35), we obtain

$$\left\| R\left(\widetilde{X} \right) \right\| - (1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right)$$
$$\leq \left(\frac{(1 - \Sigma)^2}{1 + \Sigma + 2\sqrt{\Sigma}} - (1 - \Sigma) \right) \lambda_{\min}\left(\widetilde{X} \right) \quad (45)$$
$$\leq \frac{-2(1 - \Sigma) \lambda_{\min}\left(\widetilde{X} \right)}{1 + \Sigma} < 0,$$

which implies that

$$\theta \left\| R\left(\widetilde{X} \right) \right\| \leq \lambda_{\min}\left(\widetilde{X} \right), \qquad \left(\lambda_{\min}\left(\widetilde{X} \right) - \theta \left\| R\left(\widetilde{X} \right) \right\| \right) I > 0. \quad (46)$$

According to Lemma 4, we obtain

$$\left\| g(\Delta X) \right\|$$
$$\leq \left(\left\| \Delta X \right\| + \frac{\left\| \Delta X \right\|^2}{\lambda_{\min}\left(\widetilde{X} \right) - \theta \left\| R\left(\widetilde{X} \right) \right\|} \right)$$
$$\times \sum_{i=1}^{m} \left\| X^{-1} A_i \right\|^2 + \left\| R\left(\widetilde{X} \right) \right\|$$
$$\leq \left(\theta \left\| R\left(\widetilde{X} \right) \right\| + \frac{\left(\theta \left\| R\left(\widetilde{X} \right) \right\| \right)^2}{\lambda_{\min}\left(\widetilde{X} \right) - \theta \left\| R\left(\widetilde{X} \right) \right\|} \right) \Sigma + \left\| R\left(\widetilde{X} \right) \right\|$$
$$= \theta \left\| R\left(\widetilde{X} \right) \right\|, \quad (47)$$

for $\Delta X \in \Psi$. That is, $g(\Psi) \subseteq \Psi$. By Brouwer fixed point theorem, there exists a $\Delta X \in \Psi$ such that $g(\Delta X) = \Delta X$. Hence $\widetilde{X} + \Delta X$ is a solution of (1). Moreover, by Lemma 1, we know that the solution X of (1) is unique. Then

$$\left\| \widetilde{X} - X \right\| = \left\| \Delta X \right\| \leq \theta \left\| R\left(\widetilde{X} \right) \right\|. \quad (48)$$
□

5. Perturbation Bounds

In this section we develop a relative perturbation bound for the unique solution of (1), which does not need any knowledge of the actual solution X of (1) and is easy to calculate.

Here we consider the perturbed equation:

$$\widetilde{X} - \sum_{i=1}^{m} \widetilde{A_i}^* \widetilde{X}^{-1} \widetilde{A_i} = \widetilde{Q}, \quad (49)$$

where $\widetilde{A_i}$ and $\widetilde{Q}$ are small perturbations of A_i and Q in (1), respectively. It follows from Lemma 1 that the solutions of (1) and (49) exist. Then we assume that X and $\widetilde{X}$ are the solutions of (1) and (49), respectively. Let $\Delta X = \widetilde{X} - X$, $\Delta Q = \widetilde{Q} - Q$, and $\Delta A_i = \widetilde{A_i} - A_i$.

Theorem 6. *If*

$$\sum_{i=1}^{m} \left\| A_i \right\|^2 < \beta^2, \quad (50)$$

then

$$\frac{\left\| \Delta X \right\|}{\left\| X \right\|} \leq \frac{\sum_{i=1}^{m} \left(2 \left\| A_i \right\| + \left\| \Delta A_i \right\| \right) \left\| \Delta A_i \right\| + \left\| \Delta Q \right\|}{\beta^2 - \sum_{i=1}^{m} \left\| A_i \right\|^2} \triangleq \xi_1. \quad (51)$$

Proof. By Lemma 1, we know that X and $\widetilde{X}$ are the unique solutions to (1) and (49), respectively. Subtracting (1) from (49) we have

$$\Delta X + \sum_{i=1}^{m} A_i^* \widetilde{X}^{-1} \Delta X X^{-1} A_i$$

$$= \sum_{i=1}^{m} \left(A_i^* \widetilde{X}^{-1} \Delta A_i + \Delta A_i^* \widetilde{X}^{-1} A_i \right. \tag{52}$$

$$\left. + \Delta A_i^* \widetilde{X}^{-1} \Delta A_i \right) + \Delta Q.$$

Then

$$\left\| \Delta X + \sum_{i=1}^{m} A_i^* \widetilde{X}^{-1} \Delta X X^{-1} A_i \right\|$$

$$\geq \|\Delta X\| - \left\| \sum_{i=1}^{m} A_i^* \widetilde{X}^{-1} \Delta X X^{-1} A_i \right\| \tag{53}$$

$$\geq \left(1 - \sum_{i=1}^{m} \|A_i\|^2 \left\| \widetilde{X}^{-1} \right\| \left\| X^{-1} \right\| \right) \|\Delta X\|.$$

By Lemma 2, it follows that $\|X^{-1}\| \leq 1/\beta$ and $\|\widetilde{X}^{-1}\| \leq 1/\beta$. Then

$$\left\| \Delta X + \sum_{i=1}^{m} A_i^* \widetilde{X}^{-1} \Delta X X^{-1} A_i \right\| \geq \frac{1}{\beta^2} \left(\beta^2 - \sum_{i=1}^{m} \|A_i\|^2 \right) \|\Delta X\|. \tag{54}$$

Condition (50) ensures $(\beta^2 - \sum_{i=1}^{m} \|A_i\|^2)\|\Delta X\| > 0$.

Combining (52) and (54), we obtain

$$\frac{\|\Delta X\|}{\|X\|} \leq \frac{\sum_{i=1}^{m} \left(2\|A_i\| + \|\Delta A_i\| \right) \|\Delta A_i\| + \|\Delta Q\|}{\beta^2 - \sum_{i=1}^{m} \|A_i\|^2}. \tag{55}$$

$\square$

Remark 7. Yin and Fang [20] obtained two perturbation bounds, which were dependent on the exact solution of (1), whereas, in this paper, the relative perturbation bound in Theorem 6 does not need any knowledge of the actual solution X of (1), which is important in many practical calculations.

Remark 8. With

$$\xi_1 = \frac{\sum_{i=1}^{m} \left(2\|A_i + \|\Delta A_i\|\| \right) \|\Delta A_i\| + \|\Delta Q\|}{\beta^2 - \sum_{i=1}^{m} \|A_i\|^2}, \tag{56}$$

we get $\xi_1 \to 0$ as $\|\Delta Q\| \to 0$ and $\|\Delta A_i\| \to 0$ ($i = 1, 2, \ldots, m$). Therefore (1) is well-posed.

6. Numerical Examples

To illustrate the theoretical results of the previous sections, in this section several simple examples are given, which were carried out using MATLAB 7.1. For the stopping criterion we take $\varepsilon_{k+1}(X) = \|X_k - \sum_{i=1}^{m} A_i^* X_k^{-1} A_i - Q\| < 1.0e - 10$.

TABLE 1: Backward error for Example 1 with different values of j.

j	$\|\widetilde{X} - X\|_F$	r	$U(r)$	$B(r)$
1	0.2298	0.0633	0.0630	0.0637
3	2.3×10^{-3}	6.3391×10^{-4}	6.3387×10^{-4}	6.3394×10^{-4}
5	2.2978×10^{-5}	6.3391×10^{-6}	6.3391×10^{-6}	6.3391×10^{-6}
7	2.2978×10^{-7}	6.3391×10^{-8}	6.3391×10^{-8}	6.3391×10^{-8}
9	2.2978×10^{-9}	6.3391×10^{-10}	6.3391×10^{-10}	6.3391×10^{-10}

Example 1. In this example, we consider the backward error of an approximate solution for the unique solution X to (1) in Theorem 3. We consider

$$X - A_1^* X^{-1} A_1 - A_2^* X^{-1} A_2 = Q, \tag{57}$$

with the coefficient matrices

$$A_1 = \frac{1}{5} \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}, \qquad A_2 = \frac{2\sqrt{3}}{45} A_1, \tag{58}$$

$$Q = X - A_1^* X^{-1} A_1 - A_2^* X^{-1} A_2,$$

where $X = \mathrm{diag}(1, 2, 3)$.

Let

$$\widetilde{X} = X + \begin{pmatrix} 0.5 & -0.1 & 0.2 \\ -0.1 & 0.3 & 0.6 \\ 0.2 & 0.6 & -0.4 \end{pmatrix} \times 10^{-j} \tag{59}$$

be an approximate solution to (1). Take $\alpha_1 = \|A_1\|_F$, $\alpha_2 = \|A_2\|_F$ and $\rho = \|Q\|_F$ in Theorem 3. Some results on lower and upper bounds for the backward error $\eta(\widetilde{X})$ are displayed in Table 1.

The results listed in Table 1 show that the backward error of $\widetilde{X}$ decreases as the error $\|\widetilde{X} - X\|_F$ decreases.

Example 2. This example considers the residual bound of an approximate solution for the unique solution X to (1) in Theorem 5. We consider

$$X - A_1^* X^{-1} A_1 - A_2^* X^{-1} A_2 = Q, \tag{60}$$

with

$$A_1 = \frac{(1/3) + 2 \times 10^{-2}}{\|A\|} A, \qquad A_2 = \frac{(1/6) + 3 \times 10^{-2}}{\|A\|} A,$$

$$Q = A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}. \tag{61}$$

Choose $\widetilde{X}_0 = A$. Let the approximate solution $\widetilde{X}_k$ of X be given with the iterative method $X_k = Q + \sum_{i=1}^{m} A_i^* X_{k-1}^{-1} A_i$, $X_0 > 0$, $k = 1, 2, \ldots$, where k is the iteration number.

TABLE 2: Residual bounds for Example 2 with different values of k.

k	1	2	3	4
$\left\|\widetilde{X}_k - X\right\|$	5.0268×10^{-4}	5.7662×10^{-6}	6.6162×10^{-8}	7.5024×10^{-10}
$\theta\left\|R(\widetilde{X}_k)\right\|$	5.1435×10^{-4}	5.9000×10^{-6}	6.7689×10^{-8}	7.7656×10^{-10}

TABLE 3: Perturbation bounds for Example 3 with different values of j.

j	4	5	6	7
$\left\|\widetilde{X} - X\right\| / \|X\|$	2.7482×10^{-5}	2.4983×10^{-6}	2.5705×10^{-7}	2.9406×10^{-8}
ξ_1	1.0845×10^{-4}	9.2695×10^{-6}	9.6710×10^{-7}	1.0595×10^{-7}

The residual $R(\widetilde{X}_k) \equiv Q + A_1^* \widetilde{X}_k^{-1} A_1 + A_2^* \widetilde{X}_k^{-1} A_2 - \widetilde{X}_k$ satisfies the conditions in Theorem 5. By Theorem 5, we can compute the residual bounds for $\widetilde{X}_k$

$$\left\|\widetilde{X}_k - X\right\| \le \theta \left\|R\left(\widetilde{X}_k\right)\right\|, \tag{62}$$

where

$$
\begin{aligned}
\theta = &\left(2\lambda_{\min}\left(\widetilde{X}_k\right)\right) \\
&\times \left((1 - \Sigma)\,\lambda_{\min}\left(\widetilde{X}_k\right) + \left\|R\left(\widetilde{X}_k\right)\right\|\right. \\
&\quad + \left(\left(\left((1 - \Sigma)\,\lambda_{\min}\left(\widetilde{X}_k\right) + \left\|R\left(\widetilde{X}_k\right)\right\|\right)^2\right.\right. \\
&\quad\quad \left.\left.\left. -4\lambda_{\min}\left(\widetilde{X}_k\right)\left\|R\left(\widetilde{X}_k\right)\right\|\right)^{1/2}\right)^{-1}.
\end{aligned}
\tag{63}
$$

Some results are listed in Table 2.

The results listed in Table 2 show that the residual bound given by Theorem 5 is fairly sharp.

Example 3. In this example, we consider the corresponding perturbation bound for the solution X in Theorem 6.

We consider the matrix equation

$$X - A_1^* X^{-1} A_1 - A_2^* X^{-1} A_2 = I, \tag{64}$$

with

$$A_1 = \frac{(1/3) + 2 \times 10^{-2}}{\|A\|} A, \qquad A_2 = \frac{(1/6) + 3 \times 10^{-2}}{\|A\|} A, \tag{65}$$

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}. \tag{66}$$

Suppose that the coefficient matrices A_1 and A_2 are perturbed to $\widetilde{A}_i = A_i + \Delta A_i$, $i = 1, 2$, where

$$
\begin{aligned}
\Delta A_1 &= \frac{10^{-j}}{\|C^T + C\|}\left(C^T + C\right), \\
\Delta A_2 &= \frac{3 \times 10^{-j-1}}{\|C^T + C\|}\left(C^T + C\right)
\end{aligned}
\tag{67}
$$

and C is a random matrix generated by MATLAB function randn.

By Theorem 6, we can compute the relative perturbation bound ξ_1. The results averaged as the geometric mean of 20 randomly perturbed runs. Some results are listed in Table 3.

The results listed in Table 3 show that the perturbation bound ξ_1 given by Theorem 6 is fairly sharp.

7. Concluding Remarks

In this paper, we consider the sensitivity analysis of the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$. Compared with existing literature, the contributions of this paper are as follows.

(i) A backward error and a computable residual bound of an approximate solution for the unique solution to (1) are derived, which do not appear in other known literature works.

(ii) Some results in this paper can cover the work of Li and Zhang [13] for the matrix equation $X - A^* X^{-1} A = Q$ as a special case.

(iii) This paper develops a new relative perturbation bound for the solution to (1), which does not need any knowledge of the actual solution X of (1) and could be computed easily.

Acknowledgments

References

[1] W. N. Anderson, G. B. Kleindorfer, M. B. Kleindorfer, and M. B. Woodroofe, "Consistent estimates of the parameters of a linear systems," *The Annals of Mathematical Statistics*, vol. 40, no. 6, pp. 2064–2075, 1969.

[2] W. N. Anderson, T. D. Morley, and G. E. Trapp, "The cascade limit, the shorted operator and quadratic optimal control,"

in *Linear Circuits, Systems and Signal Processsing: Theory and Application*, I. Christopher Byrnes, F. C. Martin, and E. Richard Saeks, Eds., pp. 3–7, North-Holland, New York, NY, USA, 1988.

[3] R. S. Bucy, "A priori bounds for the Riccati equation," in *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Probability Theory*, pp. 645–656, University of California Press, Berkeley, Calif, USA, 1972.

[4] D. V. Ouellette, "Schur complements and statistics," *Linear Algebra and Its Applications C*, vol. 36, pp. 187–295, 1981.

[5] W. Pusz and S. L. Woronowitz, "Functional caculus for sequlinear forms and purification map," *Reports on Mathematical Physics*, vol. 8, no. 2, pp. 159–170, 1975.

[6] J. Zabcyk, "Remarks on the control of discrete time distributed parameter systems," *SIAM Journal on Control and Optimization*, vol. 12, no. 4, pp. 721–735, 1974.

[7] A. C. M. Ran and M. C. B. Reurings, "A nonlinear matrix equation connected to interpolation theory," *Linear Algebra and Its Applications*, vol. 379, no. 1–3, pp. 289–302, 2004.

[8] X. F. Duan and A. P. Liao, "On Hermitian positive definite solution of the matrix equation $X - \sum_{i=1}^{m} A_i^* X^r A_i = Q$," *Journal of Computational and Applied Mathematics*, vol. 229, no. 1, pp. 27–36, 2009.

[9] A. Ferrante and B. C. Levy, "Hermitian solutions of the equation $X = Q + N X^{-1} N^*$," *Linear Algebra and Its Applications*, vol. 247, pp. 359–373, 1996.

[10] C. Guo and P. Lancaster, "Iterative solution of two matrix equations," *Mathematics of Computation*, vol. 68, no. 228, pp. 1589–1603, 1999.

[11] V. I. Hasanov, "Positive definite solutions of the matrix equations $X \pm A^* X^{-q} A = Q$," *Linear Algebra and Its Applications*, vol. 404, no. 1–3, pp. 166–182, 2005.

[12] V. I. Hasanov, I. G. Ivanov, and F. Uhlig, "Improved perturbation estimates for the matrix equations $X \pm A^* X^{-1} A = Q$," *Linear Algebra and Its Applications*, vol. 379, no. 1–3, pp. 113–135, 2004.

[13] J. Li and Y. H. Zhang, "The Hermitian positive definite solutions and perturbation analysis of the matrix equation $X - A^* X^{-1} A = Q$," *Bulletin of the Institute of Mathematics Academia Sinica*, vol. 30, pp. 129–142, 2008 (Chinese).

[14] Y. Lim, "Solving the nonlinear matrix equation $X = Q + \sum_{i=1}^{m} M_i X^{\delta_i} M_i^*$ via a contraction principle," *Linear Algebra and Its Applications*, vol. 430, no. 4, pp. 1380–1383, 2009.

[15] V. I. Hasanov and I. G. Ivanov, "On two perturbation estimates of the extreme solutions to the equations $X \pm A^* X^{-1} A = Q$," *Linear Algebra and Its Applications*, vol. 413, no. 1, pp. 81–92, 2006.

[16] J. Li and Y. Zhang, "Perturbation analysis of the matrix equation $X - A^* X^{-q} A = Q$," *Linear Algebra and Its Applications*, vol. 431, no. 9, pp. 1489–1501, 2009.

[17] X. G. Liu and H. Gao, "On the positive definite solutions of the matrix equations $X^s \pm A^T X^{-t} A = I_n$," *Linear Algebra and Its Applications*, vol. 368, pp. 83–87, 2003.

[18] 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 and Its Applications*, vol. 362, pp. 211–228, 2003.

[19] S. F. Xu, "Perturbation analysis of the maximal solution of the matrix equation $X + A^* X^{-1} A = P$," *Linear Algebra and Its Applications*, vol. 336, no. 1–3, pp. 61–70, 2001.

[20] X. Y. Yin and L. Fang, "Perturbation analysis for the positive definite solution of the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$," *Journal of Applied Mathematics and Computing*, vol. 43, no. 1, pp. 199–211, 2013.