



On matrix equations $X \pm A^*X^{-2}A = I$

I.G. Ivanov*, V.I. Hasanov, B.V. Minchev

Faculty of Mathematics and Informatics, Shoumen University, Shoumen 9712, Bulgaria

Received 7 March 1999; accepted 11 September 2000

Submitted by F. Uhlig

Abstract

The two matrix equations $X + A^*X^{-2}A = I$ and $X - A^*X^{-2}A = I$ are studied. We construct iterative methods for obtaining positive definite solutions of these equations. Sufficient conditions for the existence of two different solutions of the equation $X + A^*X^{-2}A = I$ are derived. Sufficient conditions for the existence of positive definite solutions of the equation $X - A^*X^{-2}A = I$ are given. Numerical experiments are discussed. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 65F10

Keywords: Matrix equation; Positive definite solution; Iterative method

1. Introduction

We consider the two matrix equations

$$X + A^*X^{-2}A = I \quad (1)$$

and

$$X - A^*X^{-2}A = I, \quad (2)$$

where I is the $n \times n$ unit matrix and A is an $n \times n$ invertible matrix. Several authors [2–5] have studied the matrix equation $X \pm A^*X^{-1}A = I$ and they have obtained theoretical properties of these equations. Nonlinear matrix equations of type (1) and (2) arise in dynamic programming, stochastic filtering, control theory and statistics [4].

* Corresponding author.

Email address: i.gantchev@fmi.shu-bg.net (I.G. Ivanov), vejdi@dir.bg (V.I. Hasanov).

In many applications one must solve a system of linear equations [1] $Mx = f$, where the positive definite matrix M arises from a finite difference approximation to an elliptic partial differential operator. As an example, let

$$M = \begin{pmatrix} I & A \\ A^* & I \end{pmatrix}.$$

Solving this linear system leads us to computing a solution of one of the equations $X + A^*X^{-1}A = I$, $X + A^*X^{-2}A = I$, or $X - A^*X^{-2}A = I$. This approach was considered in [6]. The system $Mx = f$ can be solved by Woodbury's formula [7].

In this paper we show that there are two positive definite solutions of Eq. (1). We study the existence question and how to find a Hermitian positive definite solution X of (2). We propose two iterative methods which converge to a positive definite solution of (2). When norm of A is small enough the considered iteration methods converge. The rate of convergence of these methods depends on two parameters. Numerical examples are discussed and some results of the experiments are given.

We start with some notations which we use throughout this paper. We shall use $\|A\|$ to denote the spectral norm of the matrix A , i.e., $\|A\| = \sqrt{\max_i \lambda_i}$, where the λ_i are the eigenvalues of AA^* . Let matrices P and Q be Hermitian. The notation $P > Q$ ($P \geq Q$) means that $P - Q$ is positive definite (semidefinite). The assumption $P \geq Q > 0$ implies $P^{-1} \leq Q^{-1}$ and $\sqrt{P} \geq \sqrt{Q}$.

2. The matrix equation $X + A^*X^{-2}A = I$

In this section we generalize theorems which were proved in [6]. We shall assume that $\|A\| \leq \frac{2}{\sqrt{27}}$. In [6] two iterative methods converging to a positive definite solution were investigated.

We consider the matrix sequence

$$X_0 = \gamma I, \quad X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*}, \quad k = 0, 1, 2, \dots \quad (3)$$

For this iterative method we have proved [6, Theorem 2].

Theorem 1. *If there exist the numbers α, β with $0 < \alpha < \beta < 1$ for which the inequalities*

$$\alpha^2(1 - \alpha)I < AA^* < \beta^2(1 - \beta)I$$

are satisfied, then Eq. (1) has a positive definite solution.

We can change this theorem slightly. Consider the scalar function $\varphi(x) = x^2(1 - x)$. We have

$$\max_{[0,1]} \varphi(x) = \varphi\left(\frac{2}{3}\right).$$

This function is monotonically increasing where $x \in (0, \frac{2}{3})$ and monotonically decreasing where $x \in (\frac{2}{3}, 1)$. We obtain the following.

Theorem 2. *If there exist numbers α, β with $0 < \alpha \leq \beta \leq \frac{2}{3}$ for which the inequalities*

$$\alpha^2(1 - \alpha)I \leq AA^* \leq \beta^2(1 - \beta)I$$

are satisfied, then Eq. (1) has a positive definite solution X' such that $\alpha I \leq X' \leq \beta I$.

We are proving the new theorem:

Theorem 3. *Let $\tilde{\alpha}$ and $\tilde{\beta}$ be solutions of scalar equations $\tilde{\alpha}^2(1 - \tilde{\alpha}) = \min \lambda_i(AA^*)$ and $\tilde{\beta}^2(1 - \tilde{\beta}) = \max \lambda_i(AA^*)$, respectively. Assume $0 < \tilde{\alpha} \leq \tilde{\beta} \leq \frac{2}{3}$. Consider $\{X_k\}$ defined by (3). Then:*

- (i) *If $\gamma \in [0, \tilde{\alpha}]$, then $\{X_k\}$ is monotonically increasing and converges to a positive definite solution $X_{\tilde{\alpha}}$.*
- (ii) *If $\gamma \in [\tilde{\beta}, \frac{2}{3}]$, then $\{X_k\}$ is monotonically decreasing and converges to a positive definite solution $X_{\tilde{\beta}}$.*
- (iii) *If $\gamma \in (\tilde{\alpha}, \tilde{\beta})$ and $(\tilde{\beta}^2/2\tilde{\alpha}(1 - \tilde{\beta})) < 1$, then $\{X_k\}$ converges to a positive definite solution X_γ .*

Proof. Since the function $\varphi(x) = x^2(1 - x)$ is monotonically increasing where $x \in [0, \frac{2}{3}]$ we have $0 < \alpha \leq \tilde{\alpha} \leq \tilde{\beta} \leq \beta \leq \frac{2}{3}$ and the inequalities

$$\alpha^2(1 - \alpha)I \leq AA^* \leq \beta^2(1 - \beta)I$$

are satisfied.

- (i) Assume $\gamma \in [0, \tilde{\alpha}]$. Hence $X_0 = \gamma I \leq \tilde{\beta} I$. We have

$$X_1 = \sqrt{A(I - \gamma I)^{-1}A^*} \geq \sqrt{\frac{\tilde{\alpha}^2(1 - \tilde{\alpha})}{1 - \gamma}}I \geq X_0,$$

$$X_1 = \sqrt{A(I - \gamma I)^{-1}A^*} \leq \sqrt{\frac{\tilde{\beta}^2(1 - \tilde{\beta})}{1 - \gamma}}I \leq \tilde{\beta} I.$$

Thus $X_0 \leq X_1 \leq \tilde{\beta} I$. Assume that $X_{k-1} \leq X_k \leq \tilde{\beta} I$. For X_{k+1} compute

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} \geq \sqrt{A(I - X_{k-1})^{-1}A^*} = X_k,$$

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} \leq \sqrt{A(I - \tilde{\beta} I)^{-1}A^*} \leq \sqrt{\frac{\tilde{\beta}^2(1 - \tilde{\beta})}{1 - \tilde{\beta}}}I = \tilde{\beta} I.$$

Hence $\{X_k\}$ is monotonically increasing and bounded from above by the matrix $\tilde{\beta} I$. Consequently the sequence $\{X_k\}$ converges to a positive definite solution $X_{\tilde{\alpha}}$.

(ii) Assume $\gamma \in [\tilde{\beta}, \frac{2}{3}]$. Hence $X_0 = \gamma I \geq \tilde{\alpha} I$. We have

$$X_1 = \sqrt{A(I - \gamma I)^{-1}A^*} \leq \sqrt{\frac{\tilde{\beta}^2(1 - \tilde{\beta})}{1 - \gamma}} I \leq \gamma I,$$

$$X_1 = \sqrt{A(I - \gamma I)^{-1}A^*} \geq \sqrt{\frac{\tilde{\alpha}^2(1 - \tilde{\alpha})}{1 - \gamma}} I \geq \tilde{\alpha} I.$$

Thus $X_0 \geq X_1 \geq \tilde{\alpha} I$. Assume that $X_{k-1} \geq X_k \geq \tilde{\alpha} I$. For X_{k+1} compute

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} \leq \sqrt{A(I - X_{k-1})^{-1}A^*} = X_k,$$

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} \geq \sqrt{A(I - \tilde{\alpha} I)^{-1}A^*} \geq \sqrt{\frac{\tilde{\alpha}^2(1 - \tilde{\alpha})}{1 - \tilde{\alpha}}} I = \tilde{\alpha} I.$$

Hence $\{X_k\}$ is monotonically decreasing and bounded from below by the matrix $\tilde{\alpha} I$. Consequently the sequence $\{X_k\}$ converges to a positive definite solution $X_{\tilde{\beta}}$.

(iii) Assume $\gamma \in (\tilde{\alpha}, \tilde{\beta})$. Hence $\tilde{\alpha} I < X_0 = \gamma I < \tilde{\beta} I$. We prove that $\{X_k\}$ is a Cauchy sequence.

Assume that $\tilde{\alpha} I < X_k < \tilde{\beta} I$. For X_{k+1} compute

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} < \sqrt{\frac{AA^*}{1 - \tilde{\beta}}} \leq \sqrt{\frac{\tilde{\beta}^2(1 - \tilde{\beta})}{1 - \tilde{\beta}}} I = \tilde{\beta} I,$$

$$X_{k+1} = \sqrt{A(I - X_k)^{-1}A^*} > \sqrt{\frac{AA^*}{1 - \tilde{\alpha}}} \geq \sqrt{\frac{\tilde{\alpha}^2(1 - \tilde{\alpha})}{1 - \tilde{\alpha}}} I = \tilde{\alpha} I.$$

Hence $\tilde{\alpha} I < X_k < \tilde{\beta} I$ for $k = 0, 1, \dots$

Let us consider the norm

$$\|X_{k+p} - X_k\| = \left\| \sqrt{A(I - X_{k+p-1})^{-1}A^*} - \sqrt{A(I - X_{k-1})^{-1}A^*} \right\|.$$

We denote

$$P = A(I - X_{k+p-1})^{-1}A^* \quad \text{and} \quad Q = \sqrt{A(I - X_{k-1})^{-1}A^*}$$

and use the equality

$$\sqrt{P} (\sqrt{P} - \sqrt{Q}) + (\sqrt{P} - \sqrt{Q}) \sqrt{Q} = P - Q.$$

Obviously, $Y = \sqrt{P} - \sqrt{Q}$ is a solution of the linear matrix equation $\sqrt{P}Y + Y\sqrt{Q} = P - Q$. According to [8, Theorem 8.5.2] we have

$$Y = \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt. \quad (4)$$

Thus

$$\begin{aligned}\|X_{k+p} - X_k\| &= \left\| \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt \right\| \\ &\leq \int_0^\infty \|P - Q\| \|e^{-\sqrt{P}t}\| \|e^{-\sqrt{Q}t}\| dt.\end{aligned}$$

Since $\tilde{\alpha}I < X_k$ for $k = 0, 1, \dots$, $\sqrt{P} = X_{k+p}$, $\sqrt{Q} = X_k$, then $\sqrt{P} > \tilde{\alpha}I$ and $\sqrt{Q} > \tilde{\alpha}I$. Hence

$$\begin{aligned}\|X_{k+p} - X_k\| &< \int_0^\infty e^{-2\tilde{\alpha}t} dt \|P - Q\| \\ &= \frac{1}{2\tilde{\alpha}} \|A[(I - X_{k+p-1})^{-1} - (I - X_{k-1})^{-1}]A^*\| \\ &= \frac{1}{2\tilde{\alpha}} \|A(I - X_{k-1})^{-1}(X_{k+p-1} - X_{k-1})(I - X_{k+p-1})^{-1}A^*\| \\ &\leq \frac{1}{2\tilde{\alpha}} \|A\|^2 \|(I - X_{k-1})^{-1}\| \|(I - X_{k+p-1})^{-1}\| \\ &\quad \|X_{k+p-1} - X_{k-1}\| \\ &< \frac{\tilde{\beta}^2(1 - \tilde{\beta})}{2\tilde{\alpha}} \frac{1}{(1 - \tilde{\beta})^2} \|X_{k+p-1} - X_{k-1}\| \\ &= \left(\frac{\tilde{\beta}^2}{2\tilde{\alpha}(1 - \tilde{\beta})} \right) \|X_{k+p-1} - X_{k-1}\|.\end{aligned}$$

Consequently

$$\|X_{k+p} - X_k\| \leq \left(\frac{\tilde{\beta}^2}{2\tilde{\alpha}(1 - \tilde{\beta})} \right)^k \|X_p - X_0\|.$$

Since

$$q = \left(\frac{\tilde{\beta}^2}{2\tilde{\alpha}(1 - \tilde{\beta})} \right) < 1$$

and

$$\begin{aligned}\|X_p - X_0\| &\leq \|X_p - X_{p-1}\| + \|X_{p-1} - X_{p-2}\| + \dots + \|X_1 - X_0\| \\ &\leq (q^{p-1} + \dots + q + 1) \|X_1 - X_0\| \\ &< \frac{1}{1 - q} \|X_1 - X_0\|,\end{aligned}$$

we have

$$\|X_{k+p} - X_k\| < \frac{q^k}{1 - q} \|X_1 - X_0\|.$$

The sequence $\{X_k\}$ forms a Cauchy sequence considered in the Banach space $\mathcal{C}^{n \times n}$ (where X_k are $n \times n$ positive definite matrices). Hence this sequence has a positive definite limit of (1). \square

Theorem 4. Let α_1 and β_1 be real for which the inequalities

$$(i) \quad \alpha_1^2(1 - \alpha_1)I \leq AA^* \leq \beta_1^2(1 - \beta_1)I,$$

$$(ii) \quad \frac{\beta_1^2}{2\alpha_1(1 - \beta_1)} < 1$$

are satisfied.

For each two γ_1, γ_2 with $0 \leq \alpha_1 \leq \gamma_1 \leq \gamma_2 \leq \beta_1 \leq \frac{2}{3}$ the recurrence equation (3) defines two matrix sequences $\{X'_k\}$ and $\{X''_k\}$ with initial points $X'_0 = \gamma_1 I$ and $X''_0 = \gamma_2 I$. These sequences converge to the same limit X_γ which a positive definite solution of (1).

Proof. We have $\alpha_1 I \leq X'_k \leq \beta_1 I$ and $\alpha_1 I \leq X''_k \leq \beta_1 I$. We put $P = A(I - X'_{k-1})^{-1} A^*$, $Q = A(I - X''_{k-1})^{-1} A^*$ and for $\|X'_k - X''_k\|$ we obtain

$$\begin{aligned} \|X'_k - X''_k\| &= \left\| \sqrt{A(I - X'_{k-1})^{-1} A^*} - \sqrt{A(I - X''_{k-1})^{-1} A^*} \right\| \\ &= \left\| \sqrt{P} - \sqrt{Q} \right\| \\ &= \left\| \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt \right\| \\ &\leq \int_0^\infty e^{-\alpha_1 t} dt \left\| A \left[(I - X'_{k-1})^{-1} - (I - X''_{k-1})^{-1} \right] A^* \right\| \\ &\leq \frac{1}{2\alpha} \|A\|^2 \left\| (I - X'_{k-1})^{-1} \right\| \left\| (I - X''_{k-1})^{-1} \right\| \|X'_{k-1} - X''_{k-1}\| \\ &\leq \frac{\beta_1^2(1 - \beta_1)}{2\alpha_1} \frac{1}{(1 - \beta_1)^2} \|X'_{k-1} - X''_{k-1}\| \\ &= \frac{\beta_1^2}{2\alpha_1(1 - \beta_1)} \|X'_{k-1} - X''_{k-1}\|. \end{aligned}$$

Hence

$$\|X'_k - X''_k\| \leq \frac{\beta_1^2}{2\alpha_1(1 - \beta_1)} \|X'_{k-1} - X''_{k-1}\|.$$

But

$$\frac{\beta_1^2}{2\alpha_1(1 - \beta_1)} < 1.$$

Consequently sequences $\{X'_k\}$ and $\{X''_k\}$ have the common limit. \square

Next consider the matrix sequence

$$X_0 = \alpha I, \quad X_{s+1} = I - A^* X_s^{-2} A, \quad s = 0, 1, 2, \dots \quad (5)$$

Theorem 5. Suppose there is a matrix A and numbers η, γ for which the conditions

$$(i) \quad \frac{2}{3} < \gamma \leq \eta \leq 1,$$

$$(ii) \quad \eta^2(1 - \eta)I \leq A^*A \leq \gamma^2(1 - \gamma)I$$

are satisfied. Then $\{X_s\}$ defined by (5) for $\alpha \in [\gamma, \eta]$ converges to a positive definite solution X'' of (1) at least linearly and $\gamma I \leq X'' \leq \eta I$.

Proof. Since $X_0 = \alpha I$ we have $\gamma I \leq X_0 \leq \eta I$. Suppose $\gamma I \leq X_s \leq \eta I$. Then

$$\frac{1}{\eta^2}I \leq X_s^{-2} \leq \frac{1}{\gamma^2}I$$

and

$$\frac{1}{\eta^2}A^*A \leq A^*X_s^{-2}A \leq \frac{1}{\gamma^2}A^*A.$$

According to condition (ii) we obtain

$$(1 - \eta)I \leq \frac{1}{\eta^2}A^*A \quad \text{and} \quad \frac{1}{\gamma^2}A^*A \leq (1 - \gamma)I.$$

Hence

$$\gamma I \leq X_{s+1} = I - A^*X_s^{-2}A \leq \eta I. \quad (6)$$

We have

$$\begin{aligned} X_{s+p} - X_s &= A^* \left(X_{s-1}^{-2} - X_{s+p-1}^{-2} \right) A \\ &= A^* X_{s+p-1}^{-2} \left(X_{s+p-1}^2 - X_{s-1}^2 \right) X_{s-1}^{-2} A \\ &= A^* X_{s+p-1}^{-2} \left[X_{s+p-1} (X_{s+p-1} - X_{s-1}) \right. \\ &\quad \left. + (X_{s+p-1} - X_{s-1}) X_{s-1} \right] X_{s-1}^{-2} A \\ &= A^* X_{s+p-1}^{-1} (X_{s+p-1} - X_{s-1}) X_{s-1}^{-2} A \\ &\quad + A^* X_{s+p-1}^{-2} (X_{s+p-1} - X_{s-1}) X_{s-1}^{-1} A. \end{aligned}$$

Then

$$\begin{aligned} \|X_{s+p} - X_s\| &\leq \left(\|A^* X_{s+p-1}^{-1}\| \|X_{s-1}^{-2} A\| + \|A^* X_{s+p-1}^{-2}\| \|X_{s-1}^{-1} A\| \right) \\ &\quad \times \|X_{s+p-1} - X_{s-1}\|. \end{aligned}$$

But for all k we have $X_k^{-1} \leq \frac{1}{\gamma}I$ and thus $\|X_k^{-1}\| \leq \frac{1}{\gamma}$, $\|X_k^{-2}\| \leq \frac{1}{\gamma^2}$. Using (ii) we have $\|A\| = \sqrt{\rho(A^*A)} \leq \sqrt{\gamma^2(1 - \gamma)}$. Therefore

$$\|A^* X_k^{-1}\| \leq \|A^*\| \|X_k^{-1}\| \leq \frac{\sqrt{\gamma^2(1 - \gamma)}}{\gamma},$$

$$\|A^* X_k^{-2}\| \leq \|A^*\| \|X_k^{-2}\| \leq \frac{\sqrt{\gamma^2(1-\gamma)}}{\gamma^2}.$$

Hence for all integer $p > 0$ and arbitrary s we have

$$\|X_{s+p} - X_s\| \leq \frac{2(1-\gamma)}{\gamma} \|X_{s+p-1} - X_{s-1}\|.$$

We also have $q = 2(1-\gamma)/\gamma < 1$ since $\frac{2}{3} < \gamma$.

We obtain

$$\|X_{s+p} - X_s\| \leq q^s \|X_p - X_0\|.$$

Since

$$\|X_p - X_0\| < \frac{1}{1-q} \|X_1 - X_0\|,$$

we have

$$\|X_{s+p} - X_s\| \leq \frac{q^s}{1-q} \|X_1 - X_0\|.$$

The matrices $\{X_s\}$ form a Cauchy sequence considered in the Banach space $\mathcal{C}^{n \times n}$ (where X_s are $n \times n$ positive definite matrices). Hence this sequence has a limit X'' . Using (6) we obtain

$$\gamma I \leq X'' \leq \eta I.$$

We have

$$\begin{aligned} \|X_{s+1} - X''\| &\leq \left(\|A^* X_s^{-1}\| \|(X'')^{-2} A^*\| + \|A^* X_s^{-2}\| \|(X'')^{-1} A^*\| \right) \\ &\quad \times \|X_s - X''\|. \end{aligned}$$

Similarly,

$$\|X_{s+1} - X''\| \leq \frac{2(1-\gamma)}{\gamma} \|X_s - X''\|. \quad \square$$

Remark 1. Condition (ii) of Theorem 5 implies that $\|A\| \leq \frac{2}{\sqrt{27}}$.

Theorem 6. Suppose the iterative method (5) converges to a positive definite solution X'' . Then for any $\varepsilon > 0$

$$\|X_{s+1} - X''\| < 2 \left(\|(X'')^{-1} A\| \|(X'')^{-2} A\| + \varepsilon \right) \|X_s - X''\|$$

for all s large enough.

Proof. Since $X_s \rightarrow X''$ for any $\varepsilon > 0$, there is an integer s_0 so that for $s > s_0$

$$\|X_s^{-1} A\| \|(X'')^{-2} A\| < \|(X'')^{-1} A\| \|(X'')^{-2} A\| + \varepsilon$$

and

$$\|(X'')^{-1}A\| \|X_s^{-2}A\| < \|(X'')^{-1}A\| \|(X'')^{-2}A\| + \varepsilon.$$

Then

$$\begin{aligned} \|X_{s+1} - X''\| &\leq \left(\|A^*X_s^{-1}\| \|(X'')^{-2}A\| + \|A^*X_s^{-2}\| \|(X'')^{-1}A\| \right) \|X_s - X''\| \\ &< 2 \left(\|(X'')^{-1}A\| \|(X'')^{-2}A\| + \varepsilon \right) \|X_s - X''\|. \quad \square \end{aligned}$$

Theorem 7. If $\|A\| < \frac{2}{\sqrt{27}}$, then $\|(X'')^{-1}A\| \|(X'')^{-2}A\| < \frac{1}{2}$, where X'' is the limit of sequence (5).

Proof. From the condition $\|A\| < \frac{2}{\sqrt{27}}$ we obtain that sequence (5) is convergent to X'' , with $\frac{2}{3}I < X'' \leq I$. Then $\|(X'')^{-1}\| < \frac{3}{2}$ and $\|(X'')^{-2}\| < \frac{9}{4}$. We obtain

$$\begin{aligned} \|(X'')^{-1}A\| \|(X'')^{-2}A\| &\leq \|(X'')^{-1}\| \|(X'')^{-2}\| \|A\|^2 \\ &< \frac{3}{2} \times \frac{9}{4} \times \frac{4}{27} = \frac{1}{2}. \quad \square \end{aligned}$$

Remark 2. If $\|A\| < \frac{2}{\sqrt{27}}$, then algorithms (3) and (5) converge to two different positive definite solutions X' and X'' and $X'' > X'$ (see Example 1).

3. Solutions of the matrix equation $X - A^*X^{-2}A = I$

We will describe an iterative method which compute a positive definite solution of Eq. (2).

Consider the following sequence of matrices:

$$X_0 = \alpha I, \quad X_{s+1} = \sqrt{A(X_s - I)^{-1}A^*}, \quad s = 0, 1, 2, \dots \quad (7)$$

We have the following.

Theorem 8. If A is nonsingular and for some real $\alpha > 0$, we have

- (i) $AA^* < \alpha^2(\alpha - 1)I$,
- (ii) $\sqrt{\frac{AA^*}{\alpha - 1}} - \frac{1}{\alpha^2}A^*A > I$,
- (iii) $\frac{1}{2}\|A\|^2 \sqrt{\frac{\alpha - 1}{\mu}} \left(\frac{\sqrt{\alpha - 1}}{\sqrt{\mu} - \sqrt{\alpha - 1}} \right)^2 < 1$,

where μ is the smallest eigenvalue of the matrix AA^* , then Eq. (2) has a positive definite solution.

Proof. We consider sequence (7). According to condition (i) we have

$$X_1 = \sqrt{A(X_0 - I)^{-1}A^*} = \sqrt{\frac{AA^*}{\alpha - 1}} < \alpha I = X_0.$$

Hence $X_1 < X_0$. Applying condition (ii) we obtain

$$X_1 = \sqrt{\frac{AA^*}{\alpha - 1}} > I + \frac{1}{\alpha^2}A^*A > I, \quad I < X_1 < X_0.$$

Moreover, we have

$$(X_1 - I)^{-1} > (X_0 - I)^{-1},$$

$$X_2 = \sqrt{A(X_1 - I)^{-1}A^*} > \sqrt{A(X_0 - I)^{-1}A^*} = X_1.$$

Consequently $X_1 < X_2$. Using condition (ii) we have $X_1 - I > (1/\alpha^2)A^*A$ and

$$X_2 = \sqrt{A(X_1 - I)^{-1}A^*} < \alpha I = X_0.$$

Hence $I < X_1 < X_2 < X_0$.

Analogously one can prove that

$$I < X_1 < X_{2s+1} < X_{2s+3} < X_{2k+2} < X_{2k} < X_0 = \alpha I$$

for every positive integer s, k .

Consequently the subsequences $\{X_{2k}\}$, $\{X_{2s+1}\}$ are convergent to positive definite matrices. These sequences have a common limit. Indeed, we have

$$\|X_{2k} - X_{2k+1}\| = \left\| \sqrt{A(X_{2k-1} - I)^{-1}A^*} - \sqrt{A(X_{2k} - I)^{-1}A^*} \right\|.$$

We denote

$$P = A(X_{2k-1} - I)^{-1}A^*, \quad Q = A(X_{2k} - I)^{-1}A^*$$

and use the following equality:

$$\sqrt{P}(\sqrt{P} - \sqrt{Q}) + (\sqrt{P} - \sqrt{Q})\sqrt{Q} = P - Q.$$

Since $X_{2k-1} < X_{2k+1} < X_{2k}$ for each $k = 1, 2, \dots$, the matrix $Y = \sqrt{P} - \sqrt{Q}$ is a positive definite solution of the matrix equation

$$\sqrt{P}Y + Y\sqrt{Q} = P - Q.$$

According to [8, Theorem 8.5.2] we have

$$Y = \int_0^\infty e^{-\sqrt{P}t}(P - Q)e^{-\sqrt{Q}t} dt. \quad (8)$$

Since \sqrt{P} , \sqrt{Q} are positive definite matrices integral (8) exists and $e^{-\sqrt{P}t}(P - Q)e^{-\sqrt{Q}t} \rightarrow 0$ as $t \rightarrow \infty$. Then

$$\begin{aligned}\|X_{2k} - X_{2k+1}\| &= \left\| \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt \right\| \\ &\leq \int_0^\infty \|P - Q\| \|e^{-\sqrt{P}t}\| \|e^{-\sqrt{Q}t}\| dt.\end{aligned}$$

We have $X_1 < X_s$ for the matrices of sequence (7). Hence $X_1 - I < X_s - I$ and $(X_s - I)^{-1} < (X_1 - I)^{-1}$. We have

$$\|(X_1 - I)^{-1}\| = \frac{\sqrt{\alpha - 1}}{\sqrt{\mu} - \sqrt{\alpha - 1}}.$$

Consequently

$$\|(X_s - I)^{-1}\| < \frac{\sqrt{\alpha - 1}}{\sqrt{\mu} - \sqrt{\alpha - 1}}.$$

Furthermore

$$X_{s+1} = \sqrt{A(X_s - I)^{-1}A^*} > X_1 = \sqrt{\frac{AA^*}{\alpha - 1}} \geq \sqrt{\frac{\mu}{\alpha - 1}}I.$$

Thus, we obtain

$$\begin{aligned}\|X_{2k} - X_{2k+1}\| &\leq \|P - Q\| \int_0^\infty e^{-2\sqrt{\mu/(\alpha-1)}t} dt \\ &= \|P - Q\| \frac{1}{2} \sqrt{\frac{\alpha - 1}{\mu}} \\ &= \frac{1}{2} \sqrt{\frac{\alpha - 1}{\mu}} \|A(X_{2k-1} - I)^{-1}A^* - A(X_{2k} - I)^{-1}A^*\| \\ &= \frac{1}{2} \sqrt{\frac{\alpha - 1}{\mu}} \|A(X_{2k} - I)^{-1}(X_{2k} - X_{2k-1}) \\ &\quad \times (X_{2k-1} - I)^{-1}A^*\| \\ &\leq \frac{1}{2} \sqrt{\frac{\alpha - 1}{\mu}} \|A\|^2 \left(\frac{\sqrt{\alpha - 1}}{\sqrt{\mu} - \sqrt{\alpha - 1}} \right)^2 \|X_{2k} - X_{2k-1}\|.\end{aligned}$$

Condition (iii) now implies

$$q = \frac{1}{2} \|A\|^2 \sqrt{\frac{\alpha - 1}{\mu}} \left(\frac{\sqrt{\alpha - 1}}{\sqrt{\mu} - \sqrt{\alpha - 1}} \right)^2 < 1.$$

Consequently the subsequences $\{X_{2k}\}$, $\{X_{2s+1}\}$ are convergent and have a common positive definite limit which is a solution of the matrix equation (2). \square

Similarly we can write the following.

Theorem 9. *If there is a real $\alpha > 2$ for which*

$$(i) \quad AA^* > \alpha^2(\alpha - 1)I,$$

$$(ii) \quad \sqrt{\frac{AA^*}{\alpha - 1}} - \frac{1}{\alpha^2}A^*A < I,$$

$$(iii) \quad \frac{\|A\|^2}{2\alpha(\alpha - 1)^2} < 1,$$

then Eq. (2) has a positive definite solution.

Proof. We consider sequence (7). According to condition (i) we have

$$X_1 = \sqrt{A(X_0 - I)^{-1}A^*} = \sqrt{\frac{AA^*}{\alpha - 1}} > \alpha I = X_0.$$

Hence $X_1 > X_0$. Moreover, we have

$$(X_1 - I)^{-1} < (X_0 - I)^{-1},$$

$$X_2 = \sqrt{A(X_1 - I)^{-1}A^*} < \sqrt{A(X_0 - I)^{-1}A^*} = X_1.$$

Consequently $X_2 < X_1$. Using condition (ii) we have $X_1 - I < (1/\alpha^2)A^*A$ and

$$X_2 = \sqrt{A(X_1 - I)^{-1}A^*} > \alpha I = X_0.$$

Hence $X_0 < X_2 < X_1$.

Analogously one can prove that

$$\alpha I = X_0 < X_{2s} < X_{2s+2} < X_{2k+3} < X_{2k+1} < X_1.$$

for every positive integer s, k .

Consequently the subsequences $\{X_{2k}\}$, $\{X_{2s+1}\}$ are convergent to positive definite matrices. These sequences have a common limit. Indeed, we have

$$\|X_{2k+1} - X_{2k}\| = \left\| \sqrt{A(X_{2k} - I)^{-1}A^*} - \sqrt{A(X_{2k-1} - I)^{-1}A^*} \right\|.$$

We denote

$$P = A(X_{2k} - I)^{-1}A^*, \quad Q = A(X_{2k-1} - I)^{-1}A^*$$

and use the equality

$$\sqrt{P}(\sqrt{P} - \sqrt{Q}) + (\sqrt{P} - \sqrt{Q})\sqrt{Q} = P - Q.$$

Since $X_{2k} < X_{2k+1} < X_{2k-1}$ for each $k = 1, 2, \dots$ the matrix $Y = \sqrt{P} - \sqrt{Q}$ is a positive definite solution of the matrix equation

$$\sqrt{P} Y + Y \sqrt{Q} = P - Q.$$

According to [8, Theorem 8.5.2] we have

$$Y = \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt. \quad (9)$$

Since \sqrt{P} , \sqrt{Q} are positive definite matrices, integral (9) exists and $e^{-\sqrt{P}t}(P - Q)e^{-\sqrt{Q}t} \rightarrow 0$ as $t \rightarrow \infty$. Then

$$\begin{aligned} \|X_{2k+1} - X_{2k}\| &= \left\| \int_0^\infty e^{-\sqrt{P}t} (P - Q) e^{-\sqrt{Q}t} dt \right\| \\ &\leq \int_0^\infty \|P - Q\| e^{-\sqrt{P}t} \|e^{-\sqrt{Q}t}\| dt. \end{aligned}$$

We have $\alpha I = X_0 < X_s$ for the matrices of (7). Hence $X_0 - I < X_s - I$ and $(X_s - I)^{-1} < (X_0 - I)^{-1}$. We have $\|(X_0 - I)^{-1}\| = (1/(\alpha - 1))$. Consequently

$$\|(X_s - I)^{-1}\| < \frac{1}{\alpha - 1}, \quad \sqrt{P} = X_{2k+1} > \alpha I, \quad \sqrt{Q} = X_{2k} > \alpha I.$$

Thus we obtain

$$\begin{aligned} \|X_{2k+1} - X_{2k}\| &\leq \|P - Q\| \int_0^\infty e^{-2\alpha t} dt = \|P - Q\| \frac{1}{2\alpha} \\ &= \frac{1}{2\alpha} \|A(X_{2k} - I)^{-1} A^* - A(X_{2k-1} - I)^{-1} A^*\| \\ &= \frac{1}{2\alpha} \|A(X_{2k-1} - I)^{-1} (X_{2k-1} - X_{2k}) (X_{2k} - I)^{-1} A^*\| \\ &\leq \frac{\|A\|^2}{2\alpha(\alpha - 1)^2} \|X_{2k-1} - X_{2k}\|. \end{aligned}$$

Condition (iii) now implies

$$q = \frac{\|A\|^2}{2\alpha(\alpha - 1)^2} < 1.$$

Consequently the subsequences $\{X_{2k}\}$, $\{X_{2s+1}\}$ are convergent and have a common positive definite limit which is a solution of the matrix equation (2). \square

Consider the following matrix sequence:

$$X_0 = \eta I, \quad X_{s+1} = I + A^* X_s^{-2} A, \quad s = 0, 1, 2, \dots \quad (10)$$

Theorem 10. Suppose there is a matrix A and numbers α, β for which the conditions

- (i) $1 \leq \beta < \alpha < \frac{\beta}{2} + 1$,
- (ii) $\alpha^2(\beta - 1)I \leq A^* A \leq \beta^2(\alpha - 1)I$

are satisfied. Then $\{X_s\}$ defined by (10) for $\eta \in [\beta, \alpha]$ converges to a positive definite solution X of (2) with at least linear rate of convergence and $\beta I \leq X \leq \alpha I$.

Proof. Since $X_0 = \eta I$ we have $\beta I \leq X_0 \leq \alpha I$. Suppose $\beta I \leq X_s \leq \alpha I$. Then

$$\frac{1}{\alpha^2} A^* A \leq A^* X_s^{-2} A \leq \frac{1}{\beta^2} A^* A.$$

According to condition (ii) we obtain

$$(\beta - 1)I \leq \frac{1}{\alpha^2} A^* A \quad \text{and} \quad \frac{1}{\beta^2} A^* A \leq (\alpha - 1)I.$$

Hence

$$\beta I \leq X_{s+1} = I + A^* X_s^{-2} A \leq \alpha I. \quad (11)$$

We have

$$\begin{aligned} X_{s+p} - X_s &= A^* \left(X_{s+p-1}^{-2} - X_{s-1}^{-2} \right) A \\ &= A^* X_{s-1}^{-2} \left(X_{s-1}^2 - X_{s+p-1}^2 \right) X_{s+p-1}^{-2} A \\ &= A^* X_{s-1}^{-2} \left[X_{s-1} (X_{s-1} - X_{s+p-1}) \right. \\ &\quad \left. + (X_{s-1} - X_{s+p-1}) X_{s+p-1} \right] X_{s+p-1}^{-2} A \\ &= A^* X_{s-1}^{-1} (X_{s-1} - X_{s+p-1}) X_{s+p-1}^{-2} A \\ &\quad + A^* X_{s-1}^{-2} (X_{s-1} - X_{s+p-1}) X_{s+p-1}^{-1} A. \end{aligned}$$

Then

$$\begin{aligned} \|X_{s+p} - X_s\| &\leq \left(\|A^* X_{s-1}^{-1}\| \|X_{s+p-1}^{-2} A\| + \|A^* X_{s-1}^{-2}\| \|X_{s+p-1}^{-1} A\| \right) \\ &\quad \times \|X_{s-1} - X_{s+p-1}\|. \end{aligned}$$

But for all k we have $\|X_k^{-1}\| \leq 1/\beta$ and $\|X_k^{-2}\| \leq 1/\beta^2$. Using (ii) we have $\|A\| \leq \sqrt{\beta^2(\alpha - 1)}$. Therefore

$$\|A^* X_k^{-1}\| \leq \frac{\sqrt{\beta^2(\alpha - 1)}}{\beta}, \quad \|A^* X_k^{-2}\| \leq \frac{\sqrt{\beta^2(\alpha - 1)}}{\beta^2}.$$

Hence for all integer $p > 0$ and arbitrary s we have

$$\|X_{s+p} - X_s\| \leq \frac{2(\alpha - 1)}{\beta} \|X_{s+p-1} - X_{s-1}\|.$$

Using (i) $\alpha < \frac{\beta}{2} + 1$ we have

$$q = \frac{2(\alpha - 1)}{\beta} < 1.$$

We obtain

$$\|X_{s+p} - X_s\| \leq q^s \|X_p - X_0\|.$$

Since

$$\|X_p - X_0\| < \frac{1}{1-q} \|X_1 - X_0\|,$$

we have

$$\|X_{s+p} - X_s\| < \frac{q^s}{1-q} \|X_1 - X_0\|.$$

The matrix sequence $\{X_s\}$ is a Cauchy sequence considered in the Banach space $\mathcal{C}^{n \times n}$ (where X_s are $n \times n$ positive definite matrices). Hence this sequence has a limit X . From (11) we obtain

$$\beta I \leq X \leq \alpha I.$$

We have

$$\|X_{s+1} - X\| \leq \left(\|A^* X^{-1}\| \|X_s^{-2} A^*\| + \|A^* X^{-2}\| \|X_s^{-1} A^*\| \right) \|X_s - X\|.$$

Similarly,

$$\|X_{s+1} - X\| \leq \frac{2(\alpha - 1)}{\beta} \|X_s - X\|. \quad \square$$

Theorem 11. Suppose the iterative method (10) converges to a positive definite solution X . Then for any $\varepsilon > 0$

$$\|X_{s+1} - X\| < 2(\|X^{-1} A\| \|X^{-2} A\| + \varepsilon) \|X_s - X\|$$

for all s large enough.

The proof is similar to that of Theorem 6.

4. Numerical experiments

We have made numerical experiments to compute a positive definite solution of Eqs. (1) and (2). The solution was computed for different matrices A and different values of n . Computations were done on a PENTIUM 200 MHz computer. All programs were written in MATLAB. We denote

$$\varepsilon_1(Z) = \|Z + A^* Z^{-2} A - I\|_\infty$$

and

$$\varepsilon_2(Z) = \|Z - A^* Z^{-2} A - I\|_\infty.$$

We have tested our iteration processes for solving the equation $X + A^*X^{-2}A = I$ on the following $n \times n$ matrices. We use the stopping criterion $\varepsilon_1(Z) < 1.0e - 8$.

Example 1. Consider Eq. (1) with

$$A = \begin{pmatrix} 0.01 & 0.02 & 0.03 & 0.04 \\ 0.01 & 0.225 & 0.12 & 0.02 \\ 0 & 0.09 & 0.07 & 0.03 \\ 0.12 & 0.01 & 0.02 & 0.19 \end{pmatrix}.$$

The spectral norm of A is 0.292.

Consider the iterative method (3). If $X_0 = \alpha I$ and $\alpha = 0$, then 15 iterations are required for computing X' . In this case we obtain a monotonically increasing sequence which converges to the solution X' .

Consider the case $X_0 = \beta I$. If $\beta = \frac{2}{3}$, then 16 iterations are required for computing X' . If $\beta = 0.368$, then 12 iterations are required for computing X' . In this case β satisfies conditions of Theorem 2. We obtain a monotonically decreasing sequence which converges to the solution X' . There is a value of β ($\beta = 0.368$) for which the iterative method (3) with $X_0 = \beta I$ is faster than the iterative method (3) with $X_0 = \alpha I$. Here β is a solution of the scalar equation $\|A\|^2 = \beta^2(1 - \beta)$.

Consider the iterative method (5). We choose initial value $X_0 = \alpha I$ and $\frac{2}{3} < \alpha \leq 1$ (see Theorem 5). In this case the matrix sequence converges to the solution X'' . If $\alpha = 1$, then 12 iterations are required for computing X'' . If $\alpha = \frac{(2/3)+1}{2}$, then 11 iterations are required for computing X'' . If $\alpha = 0.892$, then nine iterations are required for computing X'' . We choose α to be a solution of the scalar equation $\|A\|^2 = \alpha^2(1 - \alpha)$ (see Theorem 5).

Example 2. Consider Eq. (1) with

$$A = \begin{pmatrix} -0.1 & -0.1 & 0.02 & 0.08 \\ -0.09 & 0.3 & -0.2 & -0.1 \\ -0.04 & 0.1 & 0.01 & -0.1 \\ -0.08 & -0.06 & -0.1 & -0.2 \end{pmatrix}.$$

For this example the conditions of Theorems 2 and 5 are not satisfied ($\|A\| = 0.422 > \sqrt{\frac{4}{27}}$). But the iterative method (5) is convergent. We use this method for computing the X'' . If $\eta = \frac{2}{3}$, then 14 iterations are required for computing X'' . If $\eta = 1$, then 13 iterations are required for computing X'' . If $\eta = \frac{(2/3)+1}{2}$, then 13 iterations are required for computing X'' .

We have tested our iteration processes for solving the equation $X - A^*X^{-2}A = I$ on the following $n \times n$ matrices. We use the stopping criterion $\varepsilon_2(Z) < 10^{-8}$.

Example 3. We have $A = U\tilde{A}U^{-1}$, where U is a randomly generated matrix and

$$\tilde{A} = \text{diag} \left[12 + \frac{1}{n}, 12 + \frac{1}{2n}, \dots, 12 + \frac{1}{n^2} \right].$$

Table 1
Iterative method (7)

n	$\ A\ $	α	k_{X_α}
5	12.2204	5.605	20
10	12.2342	5.563	22
20	12.1455	5.570	21
30	12.2832	5.524	23
50	13.2733	5.238	27

The matrix A satisfies the conditions of Theorem 9. Let k_{X_α} be the smallest number s , for which $\varepsilon_2(X_s) < 10^{-8}$ for the iterative method (7). The results are given in Table 1.

Example 4. Consider Eq. (2) with

$$A = \begin{pmatrix} 0.1 & 0.2 & -0.06 & -0.16 \\ -0.2 & -0.3 & 0.16 & 0.33 \\ 0.1 & 0 & 0.02 & 0.1 \\ 0 & 0.1 & 0 & 0.03 \end{pmatrix}.$$

Consider the iterative method (10). The matrix A satisfies the conditions of Theorem 8. We choose $\beta = 1$. We compute $\|A\| = 0.587$. If $\alpha = 1.345$, then six iterations are required for computing X .

5. Conclusion

In this paper we consider special nonlinear matrix equations. We introduce iteration algorithms by which positive definite solutions of the equations can be calculated.

It was proved in [4, Theorem 13] that if $\|A\| < \frac{1}{2}$, then the equation $X + A^*X^{-1}A = I$ has a positive definite solution. From Theorem 2 (Theorem 5) we can see that if

$$\|A\|^2 < \max \varphi(x) = x^2(1-x) = \varphi\left(\frac{2}{3}\right) = \frac{4}{27},$$

then the equation $X + A^*X^{-2}A = I$ has a positive definite solution.

There are matrices A (see Examples 3 and 4) for which $\|A\| > \frac{1}{2}$ and $\|A\| > \sqrt{\frac{4}{27}}$ and for which we can still compute a positive definite solution of the equation $X - A^*X^{-2}A = I$.

References

- [1] B.L. Buzbee, G.H. Golub, C.W. Nielson, On direct methods for solving Poisson's equations, *SIAM J. Numer. Anal.* 7 (1970) 627–656.
- [2] C. Guo, P. Lancaster, iterative solution of two matrix equations, *Math. Comput.* 68 (1999) 1589–1603.
- [3] J.C. Engwerda, A.C.M. Ran, A.L. Rijkeboer, Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^*X^{-1}A = Q$, *Linear Algebra Appl.* 186 (1993) 255–275.
- [4] J.C. Engwerda, On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1}A = I$, *Linear Algebra Appl.* 194 (1993) 91–108.
- [5] A. Ferrante, B.C. Levy, Hermitian solutions of the equation $X = Q + NX^{-1}N^*$, *Linear Algebra Appl.* 247 (1996) 359–373.
- [6] I.G. Ivanov, S.M. El-Sayed, Properties of positive definite solution of the equation $X + A^*X^{-2}A = I$, *Linear Algebra Appl.* 279 (1998) 303–316.
- [7] A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.
- [8] P. Lancaster, *Theory of Matrices*, Academic Press, New York, 1969.