

Steffensen method for solving nonlinear matrix equation $X + A^T X^{-1} A = Q$

A. Nazari^{a*}, Kh. Sayehvand^b and M. Rostami^b

^aDepartment of Mathematics, Arak University, P.O. Box 38156-8-8349, Arak, Iran.

^bFaculty of Mathematical Sciences, University of Malayer, P. O. Box 16846-13114, Malayer, Iran.

Received 8 October 2014; Revised 19 December 2014; Accepted 17 January 2015.

Abstract. In this article we study Steffensen method to solve nonlinear matrix equation $X + A^T X^{-1} A = Q$, when A is a normal matrix. We establish some conditions that generate a sequence of positive definite matrices which converges to solution of this equation.

© 2014 IAUCTB. All rights reserved.

Keywords: Fixed point method, Steffensen method, Nonlinear matrix equations.

2010 AMS Subject Classification: Primary: 50A24; Secondary: 37C25.

1. Introduction

The nonlinear matrix equation

$$X + A^T X^{-1} A = Q, \quad (1)$$

where Q is a positive definite matrix and $A \in R^{n \times n}$, has some applications in some branches of applied mathematics, for example, in optimal control theory [4, 11, 13, 14], dynamic programming [12], statistics [7, 10], and network analysis [8, 9, 15]. When Q is positive definite matrix, then there exists $Q^{-\frac{1}{2}}$, s.t. by multiplying on both side of (1) by $Q^{-\frac{1}{2}}$ we have the simple form of this equation as follows

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

*Corresponding author.

E-mail address: a-nazari@araku.ac.ir (A. Nazari).

where I is the unite matrix. There are some necessary and sufficient conditions for the existence of the solution of this equation in [4, 6]. When A is a normal matrix some necessary and sufficient conditions, for existence the positive definite solution of the matrix equation $X + A^T X^{-1} A = I$, also can be found in [4, 6]. In [4, 5] the iterative fixed point method are used for solving the matrix equation (2). In this paper, we apply the Steffensen method to solve the matrix equation (2) with initial solution $X_0^{(0)}$, when A is a normal matrix.

In section 2 we prove some of the results of fixed point method that are needed in this article.

Although when A is a normal matrix, the nonlinear matrix equation (2) was studied in [6], but in section 3 we apply the Steffensen method to solve this equation, and obtain some remarkable results for special cases of normal matrix A .

In section 4 we present some numerical examples with comparing the fixed point iterative method.

2. Some results of Fixed point method

Consider the nonlinear matrix equation (2). The following matrix sequence is a fixed point iterative method with initial value $X_0 = I$

$$\begin{aligned} X_0 &= I, \\ X_{k+1} &= I - A^T X_k^{-1} A, \quad k = 0, 1, 2, \dots \end{aligned} \quad (3)$$

Lemma 2.1 ([4], Lemma 2) If the matrix equation (2) has positive definite solution X , then $X > AA^T$.

Lemma 2.2 ([4], Lemma 4) If the sequence (3) converges, then there exists a constant $\alpha > 0$ such that $X_k > \alpha I, \forall k \in \mathbb{N}$.

Theorem 2.3 ([4], Theorem 5) The matrix sequence (3) is decreasing matrix sequence i.e. $X_{k+1} < X_k$ for $k = 0, 1, 2, \dots$

Theorem 2.4 ([4], Theorem 5) The matrix sequence (3) has a positive definite solution if and only if the matrix equation (2) has a solution.

Theorem 2.5 ([4], Theorem 11) If A is a normal matrix, then the nonlinear matrix equation (2) has a solution if and only if $\rho(A) \leq \frac{1}{2}$, where $\rho(A)$ is the spectral radiuse of A .

Lemma 2.6 If A and B are nonsingular matrices, then

- a) $AB^{-1} = B^{-1}A \Leftrightarrow AB = BA$
- b) $A^{-1}B = BA^{-1} \Leftrightarrow AB = BA$
- c) $AB = BA \Rightarrow A^2B^{-1} = B^{-1}A^2$
- d) $AB^{-1} = B^{-1}A \Rightarrow A^2B^{-1} = B^{-1}A^2$
- e) $A^2B = BA^2 \Rightarrow A^2B^{-1} = B^{-1}A^2$.

Proof. Proof is trivial. ■

Definition 2.7 Let A and B be two positive definite matrices, and let $A < B$, then X belongs to $[A, B]$ if $A \leq X$ and $X \leq B$, i.e. the matrices $X - A$ and $B - X$ are positive definite matrices.

Definition 2.8 For matrix sequence $\{X_k\}_{k=0}^{\infty}$ forward difference ΔX_k is defined as follows:

$$\Delta X_k = X_{k+1} - X_k, \quad k = 0, 1, 2, \dots,$$

and $\Delta^i X_k$ is defined by

$$\Delta^i X_k = \Delta^{i-1}(\Delta X_k), \quad i = 2, 3, 4, \dots$$

Lemma 2.9 If A is a normal matrix, and X_s be the produced sequence from iterative fixed point method (3), then

$$AX_s = X_s A, \quad s = 0, 1, 2, \dots$$

Proof. The proof is by induction on s . For $s = 0$, we have

$$AX_0 = AI = IA = X_0 A.$$

If $s = 1$, then

$$AX_1 = A(I - A^T X_0^{-1} A) = A(I - A^T A) = A - AA^T A = (I - A^T A)A = X_1 A.$$

Now assume that $AX_k = X_k A$, we have $(AX_k)^{-1} = (X_k A)^{-1}$ hence $X_k^{-1} A^{-1} = A^{-1} X_k^{-1}$,

$$AX_{k+1} = A(I - A^T X_k^{-1} A) = A - AA^T X_k^{-1} A = A - A^T A X_k^{-1} A.$$

Thus by lemma 2.6

$$AX_{k+1} = A - A^T A^2 X_k^{-1} = A - A^T X_k^{-1} A^2 = (I - A^T X_k^{-1} A)A = X_{k+1} A. \quad \blacksquare$$

Lemma 2.10 If A is a normal matrix, and X_s is the produced sequence from iterative fixed point method (3), then

$$A^T X_s = X_s A^T, \quad s = 0, 1, 2, \dots$$

Proof. The proof is similar to the proof of Lemma 2.9. \blacksquare

Lemma 2.11 If A is a normal matrix, and X_s is the produced sequence from iterative fixed point method (3), then

$$X_s X_{s+1} = X_{s+1} X_s, \quad s = 0, 1, 2, \dots$$

Proof. The proof is by induction on s . If $s = 0$, then

$$X_0 X_1 = X_0 (I - A^T X_0^{-1} A) = I - A^T A = (I - A^T X_0^{-1} A) X_0 = X_1 X_0.$$

For $s = 1$

$$\begin{aligned} X_1 X_2 &= (I - A^T X_0^{-1} A)(I - A^T X_1^{-1} A) \\ &= I - A^T X_1^{-1} A - A^T X_0^{-1} A + A^T X_0^{-1} A A^T X_1^{-1} A, \end{aligned} \quad (4)$$

and

$$\begin{aligned} X_2 X_1 &= (I - A^T X_1^{-1} A)(I - A^T X_0^{-1} A) \\ &= I - A^T X_1^{-1} A - A^T X_0^{-1} A + A^T X_1^{-1} A A^T X_0^{-1} A. \end{aligned} \quad (5)$$

To prove (4) and (5) it suffices to prove

$$A^T X_0^{-1} A A^T X_1^{-1} A = A^T X_1^{-1} A A^T X_0^{-1} A.$$

Since A is a normal matrix, $A^T A = A A^T$ and so, $A^T A A A^T = A A^T A^T A$, therefore

$$\begin{aligned} A A^T - A^T A A A^T &= A A^T - A A^T A^T A \\ \Rightarrow (I - A^T A) A A^T &= A A^T (I - A^T A) X_1 A A^T = A A^T X_1 \\ \Rightarrow A A^T X_1^{-1} &= X_1^{-1} A A^T \Rightarrow X_0^{-1} A A^T X_1^{-1} = X_1^{-1} A A^T X_0^{-1} \\ \Rightarrow A^T X_0^{-1} A A^T X_1^{-1} A &= A^T X_1^{-1} A A^T X_0^{-1} A, \end{aligned}$$

and consequently $X_0 X_1 = X_1 X_0$. Now assume that $X_{k-1} X_k = X_k X_{k-1}$, we prove that $X_k X_{k+1} = X_{k+1} X_k$. By Lemma 2.9 we have $A X_k = X_k A$, multiplying it by A^T on the right, we have $A X_k A^T = X_k A A^T$. By Lemma 2.10 we have $A A^T X_k = X_k A A^T$, if we multiply it from left and right by X_k^{-1} , then $X_k^{-1} A A^T = A A^T X_k^{-1}$. Again we multiply this latter equation from the right by X_{k-1}^{-1} to get

$$X_k^{-1} A A^T X_{k-1}^{-1} = A A^T X_k^{-1} X_{k-1}^{-1} = A A^T X_{k-1}^{-1} X_k^{-1} = X_{k-1}^{-1} A A^T X_k^{-1}.$$

Finally, by multiplying the last equations from the right by A and from the left by A^T , we have

$$A^T X_k^{-1} A A^T X_{k-1}^{-1} A = A^T X_{k-1}^{-1} A A^T X_k^{-1} A.$$

Thus

$$\begin{aligned} I - A^T X_k^{-1} A - A^T X_{k-1}^{-1} A + A^T X_k^{-1} A A^T X_{k-1}^{-1} A \\ = I - A^T X_k^{-1} A - A^T X_{k-1}^{-1} A + A^T X_{k-1}^{-1} A A^T X_k^{-1} A \\ \Rightarrow (I - A^T X_k^{-1} A) (I - A^T X_{k-1}^{-1} A) = (I - A^T X_{k-1}^{-1} A) (I - A^T X_k^{-1} A) \\ \Rightarrow X_{k+1} X_k = X_k X_{k+1}. \end{aligned}$$

■

Lemma 2.12 If A is a normal matrix, and X_s is the produced sequence from iterative fixed point method (3), then

$$X_s X_{s+2} = X_{s+2} X_s, \quad s = 0, 1, 2, \dots$$

Proof. The proof is by induction on s . If $s = 0$, then

$$X_0 X_2 = I X_2 = X_2 I = X_2 X_0.$$

for $s = 1$

$$\begin{aligned} X_1 X_3 &= (I - A^T X_0^{-1} A)(I - A^T X_2^{-1} A) \\ &= I - A^T X_2^{-1} A - A^T X_0^{-1} A + A^T X_0^{-1} A A^T X_2^{-1} A, \end{aligned} \quad (6)$$

and

$$\begin{aligned} X_3 X_1 &= (I - A^T X_2^{-1} A)(I - A^T X_0^{-1} A) \\ &= I - A^T X_2^{-1} A - A^T X_0^{-1} A + A^T X_2^{-1} A A^T X_0^{-1} A. \end{aligned} \quad (7)$$

To show (6) and (7) it is sufficient to prove $A^T X_0^{-1} A A^T X_2^{-1} A = A^T X_2^{-1} A A^T X_0^{-1} A$. By Lemma 2.9 we have $A X_0 = X_0 A \Rightarrow A X_0 A^T = X_0 A A^T$, now by Lemma 2.10, $A A^T X_0 A^T = X_0 A A^T$, if we multiply it from right and left by X_0^{-1} and by multiplying its result from the right by X_2^{-1} , so by Lemmas 2.6, 2.9, 2.10 we have

$$X_0^{-1} A A^T X_2^{-1} = X_2^{-1} A A^T X_0^{-1},$$

therefore $A^T X_0^{-1} A A^T X_2^{-1} A = A^T X_2^{-1} A A^T X_0^{-1} A$, and consequently $X_1 X_3 = X_3 X_1$.

Now assume that $X_{k-1} X_{k+1} = X_{k+1} X_{k-1}$, we prove that $X_k X_{k+2} = X_{k+2} X_k$. By Lemma 2.9 we have $A X_{k-1} = X_{k-1} A$, multiplying from the right by A^T we have $A X_{k-1} A^T = X_{k-1} A A^T$, then by Lemma 2.10, $A A^T X_{k-1} = X_{k-1} A A^T$. We multiply this equation from right and left by X_{k-1}^{-1} to get $X_{k-1}^{-1} A A^T = A A^T X_{k-1}^{-1}$. Multiplying from the right by X_{k+1}^{-1} then by hypothesis of induction and Lemmas 2.6, 2.9, 2.10, we have

$$\begin{aligned} X_{k-1}^{-1} A A^T X_{k+1}^{-1} &= X_{k+1}^{-1} A A^T X_{k-1}^{-1} \\ \Rightarrow A^T X_{k-1}^{-1} A A^T X_{k+1}^{-1} A &= A^T X_{k+1}^{-1} A A^T X_{k-1}^{-1} A. \end{aligned}$$

Therefore

$$\begin{aligned} &I - A^T X_{k-1}^{-1} A - A^T X_{k+1}^{-1} A + A^T X_{k-1}^{-1} A A^T X_{k+1}^{-1} A \\ &= I - A^T X_{k-1}^{-1} A - A^T X_{k+1}^{-1} A + A^T X_{k+1}^{-1} A A^T X_{k-1}^{-1} A \\ &\Rightarrow (I - A^T X_{k-1}^{-1} A) (I - A^T X_{k+1}^{-1} A) = (I - A^T X_{k+1}^{-1} A) (I - A^T X_{k-1}^{-1} A) \\ &\Rightarrow X_k X_{k+2} = X_{k+2} X_k \end{aligned}$$

■

Lemma 2.13 ([17]) If A and B are two symmetric positive definite of same order, and $AB = BA$ then $(AB)^{\frac{1}{2}} < \frac{A+B}{2}$.

Theorem 2.14 ([16]) If A and B are two symmetric semi-positive definite of same order, then $AB \geq 0$ if and only if $AB = BA$.

Theorem 2.15 If σ_{max} is the greatest singular value of normal matrix A , and $0 \leq \sigma_{max} \leq \frac{1}{2}$ and $\{X_k\}_{k=0}^{\infty}$ be the matrix sequence (2), then $X_k > \frac{1}{2}I \forall k \in N$.

Proof. We have $0 \leq \sigma_{max} = \sqrt{\lambda_{max}(A^T A)} \leq \frac{1}{2}$, so $0 \leq \lambda(A^T A) \leq \frac{1}{4}$ and consequently $AA^T \leq \frac{1}{4}I$, now by induction on k , we show that $X_k > \frac{1}{2}I$.

For $k = 0$, we have $X_0 = I > \frac{1}{2}I$. Let $X_k > \frac{1}{2}I$, then

$$X_{k+1} = I - A^T X_k^{-1} A = [I + A^T (X_k - AA^T)^{-1} A]^{-1}.$$

Recall that for two invertible matrices A and B we have

$$(A + B)^{-1} = A^{-1} - A^{-1}(B^{-1} + A^{-1})^{-1}A^{-1}.$$

On the other hand $A^T A \leq \frac{1}{4}I \Rightarrow \frac{1}{4}I + A^T A \leq \frac{1}{2}I < X_k$. Now we can write

$$\begin{aligned} X_{k+1} &= [I + A^T (X_k - AA^T)^{-1} A]^{-1} \\ &> I + A^T \left(\left(\frac{1}{4}I + A^T A - AA^T \right)^{-1} A \right)^{-1} \\ &= (I + 4A^T A)^{-1} \geq \frac{1}{2}I, \end{aligned}$$

where the last inequality is obtained from $I + 4A^T A \leq 2I$. ■

Corollary 2.16 If A is a normal matrix, then the matrix sequence which obtained from (2) is a decreasing positive definite sequence with lower bound $\frac{1}{2}I$. Therefore, if the equation (2) has a solution, then this solution lies in interval the $[\frac{1}{2}I, I]$.

3. Steffensen method

In this section we use the iterative Steffensen method for solving the nonlinear matrix equation $X + A^T X^{-1} A = I$, where A is a normal matrix. The Steffensen method for solving nonlinear equation is explained in [2]. Assume the sequence of fixed point iterative for solving nonlinear matrix equation $X + A^T X^{-1} A = I$ is as follows

$$\begin{aligned} X_0 &= I, \\ X_{k+1} &= I - A^T X_k^{-1} A, \quad k = 0, 1, 2, \dots \end{aligned}$$

The Δ^2 -Aietken sequence is defined by

$$\hat{X}_k = X_k - (\Delta X_k)^2 (\Delta^2 X_k)^{-1},$$

where $\Delta X_k = X_{k+1} - X_k$ and $\Delta^2 X_k = \Delta(\Delta X_k)$. By applying the Δ^2 -Aietken in fixed point iterative method we obtain the Steffensen method:

$$\begin{aligned} X_0^{(0)} &= I \\ X_1^{(k)} &= I - A^T \left(X_0^{(k)} \right)^{-1} A \\ X_2^{(k)} &= I - A^T \left(X_1^{(k)} \right)^{-1} A \\ X_0^{(k+1)} &= X_0^{(k)} - \left(\Delta X_0^{(k)} \right)^2 \left(\Delta^2 X_0^{(k)} \right)^{-1} \\ &= X_0^{(k)} - \left(X_1^{(k)} - X_0^{(k)} \right)^2 \left(X_2^{(k)} - 2X_1^{(k)} + X_0^{(k)} \right)^{-1}, \quad k = 0, 1, 2, \dots \end{aligned} \quad (8)$$

We begin this iterative method by initial approximation solution $X_0 = I$. In this section we show that the Steffensen method generates a decreasing sequence of positive definite matrices. Consequently this sequence of positive definite matrices is convergent.

Remark 1 Recall that $X_0^{(k)}$ is positive definite for $k = 1, 2, \dots$, whenever $\Delta^2 X_0^{(k)}$ is positive definite. We know that the matrix $\Delta X_0^{(k)}$ is Hermitian, so by lemma (2.17) the matrix $(\Delta X_0^{(k)})^2$ is a positive definite, furthermore by Lemma 2.14 it is necessary that the two matrices $(\Delta^2 X_0^{(k)})^{-1}$ and $(\Delta X_0^{(k)})^2$ have commutative property in ordinary product of two matrices. It is clear that the commutative property holds, since it is sufficient that $\Delta X_0^{(k)}$ and $\Delta^2 X_0^{(k)}$ have commutative property in ordinary product and this is true by Lemmas 2.9, 2.10, 2.11, 2.12 for example, for $k = 0$ we can write

$$\begin{aligned} (\Delta X_0^{(0)}) (\Delta^2 X_0^{(0)}) &= \left(X_1^{(0)} - X_0^{(0)} \right) \left(X_2^{(0)} - 2X_1^{(0)} + X_0^{(0)} \right) \\ &= X_1^{(0)} X_2^{(0)} - 2X_1^{(0)} X_1^{(0)} + X_1^{(0)} X_0^{(0)} - X_0^{(0)} X_2^{(0)} \\ &\quad + 2X_0^{(0)} X_1^{(0)} - X_0^{(0)} X_0^{(0)}. \end{aligned} \quad (9)$$

By Lemmas 2.11 and 2.12 commutativity in product holds, thus

$$\begin{aligned} &= X_2^{(0)} X_1^{(0)} - 2X_1^{(0)} X_1^{(0)} + X_0^{(0)} X_1^{(0)} - X_2^{(0)} X_0^{(0)} + 2X_1^{(0)} X_0^{(0)} - X_0^{(0)} X_0^{(0)} \\ &= \left(X_2^{(0)} - 2X_1^{(0)} + X_0^{(0)} \right) X_1^{(0)} - \left(X_2^{(0)} - 2X_1^{(0)} + X_0^{(0)} \right) X_0^{(0)} \\ &= \left(\Delta^2 X_0^{(0)} \right) \left(\Delta X_0^{(0)} \right) \end{aligned}$$

In order to we can show for, $k = 1, 2, \dots$

Theorem 3.1 Let X_k is the sequence of positive definite matrix that converg to X . Then the matrix sequence $\hat{X}_k = X_k - (\Delta x_k)^2 (\Delta^2 x_k)^{-1}$ faster than X_k converge to X , if $X_{k+1} - X = (H + T_k)(X_k - X)$ where H is matrix with $\|H\| \leq 1$ and T_k is sequence of matrices that $\lim_{k \rightarrow \infty} T_k = 0$.

Proof. Since $E_k = X_k - X$, then

$$\begin{aligned}
 X_{k+2} - 2X_{k+1} + X_k &= E_{k+2} - 2E_{k+1} + E_k \\
 &= (H + T_{k+1})(X_{k+1} - X) - 2(H + T_k)(X_k - X) + E_k \\
 &= (H + T_{k+1})(H + T_k)(E_k) - 2(H + T_k)E_k + E_k \\
 &= ((H + T_{k+1})(H + T_k) - 2(H + T_k))E_k + E_k \\
 &= ((H + T_{k+1})(H + T_k) - 2(H + T_k) + I)E_k \\
 &= (H^2 + HT_k + HT_k + T_kT_{k+1} - 2H - 2T_k + I)E_k \\
 &= \left((H - I)^2 + H(T_{k+1} + T_k) - 2T_k \right) E_k \\
 &= \left((H - I)^2 + S_k \right) E_k.
 \end{aligned}$$

Where $S_k = H(T_{k+1} + T_k) - 2T_k$, and $\lim_{k \rightarrow \infty} S_k = 0$. On the other hand we have,

$$\begin{aligned}
 X_{k+1} - X_k &= E_{k+1} - E_k \\
 &= (H + T_k)E_k - E_k \\
 &= (H + T_k - I)E_k \\
 &= ((H - I) + T_k)E_k.
 \end{aligned}$$

Therefore for the matrix sequence \hat{X}_k ,

$$\hat{X}_k - X = X_k - X - ((H - I + T_k)E_k)^2 \left(((H - I)^2 + S_k)E_k \right)^{-1}.$$

We know that $E_k = X_k - X$ and E_k is a nonsingular matrix and by Lemma 2.11, and Lemma 2.12 we have,

$$\begin{aligned}
 \hat{X}_k - X &= E_k - (H - I + T_k)^2 E_k^2 E_k^{-1} ((H - I)^2 + S_k)^{-1} \\
 &= (I - (H - I + T_k)^2 ((H - I)^2 + S_k)^{-1}) E_k.
 \end{aligned}$$

Consequently

$$\begin{aligned}
 (\hat{X}_k - X)(X_k - X)^{-1} &= (\hat{X}_k - X)E_k^{-1} \\
 &= I - (H - I + T_k)^2 ((H - I)^2 + S_k)^{-1},
 \end{aligned}$$

and

$$\lim_{k \rightarrow \infty} (\hat{X}_k - X)(X_k - X)^{-1} = \lim_{k \rightarrow \infty} \left[I - (H - I + T_k)^2 ((H - I)^2 + S_k)^{-1} \right] = 0.$$

■

Theorem 3.2 If A is a normal matrix and $(X_1^{(k)})^2 < X_0^{(k)} X_2^{(k)}$ for $k = 1, 2, \dots$ we have then the matrices $X_i^{(k)}$ that are produced by Steffensen method (8) is positive definite

for $k = 0, 1, 2, \dots$, and $i = 0, 1, 2$.

Proof. We present proof by induction on k . For $k = 0$ we have

$$\begin{aligned} X_0^{(0)} &= I > 0 \\ X_1^{(0)} &= I - A^T \left(X_0^{(0)} \right)^{-1} A = I - A^T A. \end{aligned}$$

On the other hand

$$\begin{aligned} 0 < A^T A &\leq \frac{1}{4} I \Rightarrow -\frac{1}{4} I \leq -A^T A < 0 \\ &\Rightarrow \frac{3}{4} I \leq I - A^T A < I \\ &\Rightarrow X_1^{(0)} > 0. \end{aligned}$$

For $i = 2$, $X_2^{(0)} = I - A^T \left(X_1^{(0)} \right)^{-1} A$. By Theorem 2.15 it is clear that $\frac{1}{2} I < X_1^{(0)} < I$ and so $I < \left(X_1^{(0)} \right)^{-1} < 2I$. By multiplying the above inequality by $A^T A$ we have

$$\begin{aligned} A^T A &< A^T \left(X_1^{(0)} \right)^{-1} A < 2A^T A \\ &\Rightarrow 0 < A^T \left(X_1^{(0)} \right)^{-1} A < \frac{1}{2} I \\ &\Rightarrow \frac{1}{2} I < I - A^T \left(X_1^{(0)} \right)^{-1} A < I, \end{aligned}$$

i.e. $X_2^{(0)} > 0$. In what follows we prove that $X_0^{(1)}$ is also positive definite. By Steffensen method we have

$$X_0^{(1)} = X_0^{(0)} - \left(X_1^{(0)} - X_0^{(0)} \right)^2 \left(X_2^{(0)} - 2X_1^{(0)} + X_0^{(0)} \right)^{-1}.$$

We show that $\Delta^2 X_0^{(0)} > 0$, we have

$$\begin{aligned} \Delta^2 X_0^{(0)} &= X_2^{(0)} - 2X_1^{(0)} + X_0^{(0)} \\ &= I - A^T \left(X_1^{(0)} \right)^{-1} A - 2I + 2A^T \left(X_0^{(0)} \right)^{-1} A + I \\ &= A^T \left(2I - \left(X_1^{(0)} \right)^{-1} \right) A. \end{aligned}$$

Then by Theorem 2.15 we have $\frac{1}{2}I < X_1^{(0)} < I$, so $2I > (X_1^{(0)})^{-1}$ and hence

$$\begin{aligned} 2I - (X_1^{(0)})^{-1} &> 0 \\ \Rightarrow A^T \left(2I - (X_1^{(0)})^{-1} \right) A &> 0 \\ \Rightarrow \Delta^2 X_0^{(0)} &> 0. \end{aligned}$$

Now we can write

$$A^T A \leq \frac{1}{4}I \Rightarrow (A^T A)^2 \leq \frac{1}{4}A^T A \quad (10)$$

$$\begin{aligned} -A^T A \geq -\frac{1}{4}I &\Rightarrow I - A^T A \geq \frac{3}{4}I \\ \Rightarrow (I - A^T A)^{-1} &\leq \frac{4}{3}I \\ \Rightarrow A^T (I - A^T A)^{-1} A &\leq \frac{4}{3}A^T A \leq \frac{1}{3}I \\ \Rightarrow (A^2)^T (I - A^T A)^{-1} A^2 &\leq \frac{1}{3}A^T A. \end{aligned} \quad (11)$$

Therefor by (10) and (11)

$$(A^T A)^2 + (A^2)^T (I - A^T A)^{-1} A^2 \leq \frac{7}{12}A^T A < A^T A.$$

$$\begin{aligned} 0 \leq I - 2A^T A + (A^T A)^2 &< I - A^T A - (A^2)^T (I - A^T A)^{-1} A^2 \\ &= I - A^T \left[I + A^T (I - A^T A)^{-1} A \right] A \\ &= I - A^T \left[I + (A^{-1} (I - A^T A) A^{-T})^{-1} \right] A \\ &= I - A^T \left[I + (A^{-1} A^{-T} - I)^{-1} \right] A \\ &= I - A^T [I - A^T A] A. \end{aligned}$$

Thus

$$\begin{aligned} 0 \leq (I - A^T A)^2 &< I - A^T (X_1^{(0)})^{-1} A \\ \Rightarrow 0 \leq (X_1^{(0)})^2 &< X_2^{(0)} \end{aligned}$$

$$\begin{aligned}
&\Rightarrow 0 \leq (X_1^{(0)})^2 - 2X_1^{(0)} + I < X_2^{(0)} - 2X_1^{(0)} + I \\
&\Rightarrow 0 \leq (X_1^{(0)} - I)^2 < \Delta^2 X_0^{(0)} \\
&\Rightarrow 0 \leq (\Delta X_0^{(0)})^2 < \Delta^2 X_0^{(0)} \\
&\Rightarrow (\Delta X_0^{(0)})^2 (\Delta^2 X_0^{(0)})^{-1} < I \\
&\Rightarrow I - (\Delta X_0^{(0)})^2 (\Delta^2 X_0^{(0)})^{-1} > 0 \\
&\Rightarrow X_0^{(1)} > 0.
\end{aligned}$$

Let $X_0^{(k)}$, $X_1^{(k)}$ and $X_2^{(k)}$ be positive definite matrices, we show that $X_0^{(k+1)}$ is a positive definite matrix. By hypothesis we have $(X_1^{(k)})^2 < X_0^{(k)} X_2^{(k)}$ and by Lemma 2.13 the geometric mean of two matrices is less than or equal to the arithmetic mean of the same matrices, therefore $X_1^{(k)} < \frac{X_0^{(k)} + X_2^{(k)}}{2}$. So $\Delta^2 X_0^{(k)} = X_2^{(k)} - 2X_1^{(k)} + X_0^{(k)} > 0$, on the other hand we have $(X_1^{(k)})^2 < X_0^{(k)} X_2^{(k)}$, therefore

$$\begin{aligned}
(X_1^{(k)})^2 - 2X_0^{(k)} X_1^{(k)} + (X_0^{(k)})^2 &< X_0^{(k)} X_2^{(k)} - 2X_0^{(k)} X_1^{(k)} + (X_0^{(k)})^2 \\
&\Rightarrow (X_1^{(k)} - X_0^{(k)})^2 < X_0^{(k)} (X_2^{(k)} - 2X_1^{(k)} + X_0^{(k)}) = X_0^{(k)} (\Delta^2 X_0^{(k)}).
\end{aligned}$$

We also have $\Delta^2 X_0^{(k)} > 0$, so

$$\begin{aligned}
(\Delta X_0^{(k)})^2 (\Delta^2 X_0^{(k)})^{-1} &< X_0^{(k)} \\
\Rightarrow X_0^{(k)} - (\Delta X_0^{(k)})^2 (\Delta^2 X_0^{(k)})^{-1} &> 0 \\
\Rightarrow X_0^{(k+1)} &> 0.
\end{aligned}$$

■

Theorem 3.3 If A is a normal matrix and $X_1^{(k)} < \frac{X_0^{(k)} + X_2^{(k)}}{2}$, for $k = 1, 2, \dots$ then the matrix sequence produced by Steffensen method is a decreasing matrix sequence. i.e. $X_0^{(k+1)} < X_0^{(k)}$ $k = 0, 1, 2, \dots$

Proof. We present proof by induction on k . For $k = 0$, we have

$$X_0^{(0)} - X_0^{(1)} = (\Delta X_0^{(0)})^2 (\Delta^2 X_0^{(0)})^{-1}.$$

So, it suffices to show that $\Delta^2 X_0^{(0)} > 0$.

$$\begin{aligned}
\Delta^2 X_0^{(0)} &= X_0^{(2)} - 2X_0^{(1)} + X_0^{(0)} \\
&= I - A^T (X_0^{(1)})^{-1} A - 2I + 2A^T (X_0^{(0)})^{-1} A + I \\
&= 2A^T A - A^T (X_0^{(1)})^{-1} A = A^T (2I - X_0^{(1)})^{-1} A,
\end{aligned}$$

Now, by Theorem 3.3 we see that $\frac{1}{2}I < X_0^1 < I$, then $2I - X_0^{(1)} > 0$ and so

$$A^T(2I - X_0^{(1)})^{-1}A > 0.$$

Finally $\Delta^2 X_0^{(0)} > 0$. Since $\Delta X_0^{(0)}$ is Hermitian matrix, then $(\Delta X_0^{(0)})^2$ is a positive definite matrix and since $\Delta^2 X_0^{(0)}$ and $(\Delta X_0^{(0)})^2$ have commutative property in product of two matrices, thus the matrices $(\Delta^2 X_0^{(0)})^{-1}$ and $(\Delta X_0^{(0)})^2$ have this property and so the product of these matrices is also positive definite, consequently

$$X_0^{(0)} - X_0^{(1)} > 0 \Rightarrow X_0^{(1)} < X_0^{(0)},$$

Now, assume that $X_0^{(m+1)} < X_0^{(m)}$, we show

$$X_0^{(m+2)} < X_0^{(m+1)}, \quad (12)$$

and

$$X_0^{(m+1)} - X_0^{(m+2)} = (\Delta X_0^{(m+1)})^2 (\Delta^2 X_0^{(m+1)})^{-1}.$$

By hypothesis we have $X_1^{(m+1)} < \frac{X_0^{(m+1)} + X_2^{(m+1)}}{2}$, so $\Delta^2 X_0^{(m+1)}$ is positive definite matrix and $(\Delta X_0^{(m+1)})^2$ and commutative property of product of $X_0^{(m+1)}$ and $X_0^{(m+2)}$ we have $X_0^{(m+1)} - X_0^{(m+2)} > 0$ then (12) holds. ■

Remark 2 We show that when A is a normal matrix, the iterative sequence produced by fixed point iterative method is placed in the interval $[\frac{1}{2}I, I]$, and we will expect that by applying the Steffensen method, the matrix sequence of this method is also placed in the interval $[\frac{1}{2}I, I]$. Then we prove this property for a special case of A . Let A be a normal matrix and A_i denotes the i -th row of A , which satisfies

$$A_i A_j^T = 0 \quad i < j, \quad (13)$$

The set of all normal matrices which satisfy (13) is denoted by Ω . If $A \in \Omega$, then $A^T A$ is the following diagonal matrix,

$$A^T A = \text{diag} \left(\sum_{j=1}^n (a_{1j})^2, \sum_{j=1}^n (a_{2j})^2, \dots, \sum_{j=1}^n (a_{nj})^2 \right).$$

For $i = 1, 2, \dots, n$, we have by $0 \leq A^T A \leq \frac{1}{4}I$ that $0 \leq \sum_{j=1}^n (a_{ij})^2 \leq \frac{1}{4}$ and so $X_1^{(0)}$ is a diagonal matrix. Also the matrices $X_2^{(0)}$, $X_0^{(1)}$ and ... are diagonal matrices by (8). The next Theorem shows that $X_i^{(k)}$ belong to the interval $[\frac{1}{2}I, I]$ for $k = 0, 1, 2, \dots$ and $i = 0, 1, 2$.

Theorem 3.4 If $A \in \Omega$ then the matrix sequence $X_i^{(k)}$ to the interval $[\frac{1}{2}I, I]$ for $k = 0, 1, 2, \dots$ and $i = 0, 1, 2$.

Proof. For $k = 0$ we have $(x_0^{(0)})_{ii} = 1$ for $i = 1, 2, \dots, n$, and

$$(x_1^{(0)})_{ii} = 1 - \frac{a_i}{(x_0^{(0)})_{ii}} = 1 - a_i,$$

where $a_i = \sum_{j=1}^n (a_{ij})^2$. Whereas $1 \leq a_i \leq \frac{1}{4}$, then $\frac{3}{4} \leq (x_1^{(0)})_{ii} \leq 1$ and so $X_1^{(0)} \in [\frac{1}{2}I, I]$, Also, $(x_2^{(0)})_{ii} = 1 - \frac{a_i}{(x_1^{(0)})_{ii}}$ and by (10) we have $\frac{1}{2} \leq (x_1^{(0)})_{ii} \leq 1$. Thus, $1 \leq \frac{1}{(x_1^{(0)})_{ii}} \leq 2$ and

$$\begin{aligned} a_i \leq \frac{a_i}{(x_1^{(0)})_{ii}} \leq 2a_i &\Rightarrow 0 \leq \frac{a_i}{(x_1^{(0)})_{ii}} \leq \frac{1}{2} \\ &\Rightarrow \frac{1}{2} \leq 1 - \frac{a_i}{(x_1^{(0)})_{ii}} \leq 1 \\ &\Rightarrow \frac{1}{2} \leq (x_2^{(0)})_{ii} \leq 1. \end{aligned}$$

Consequently, $X_2^{(0)} \in [\frac{1}{2}I, I]$. Now assume that for $k \geq 1$, the matrices $X_0^{(k)}, X_1^{(k)}, X_2^{(k)}$ belong to the interval $[\frac{1}{2}I, I]$. We show that $X_0^{(k+1)}$ is also belong to the interval $[\frac{1}{2}I, I]$. We know that

$$\begin{aligned} (x_0^{(k+1)})_{ii} &= (x_0^{(k)})_{ii} - \frac{((x_1^{(k)})_{ii} - (x_0^{(k)})_{ii})^2}{(x_2^{(k)})_{ii} - 2(x_1^{(k)})_{ii} + (x_0^{(k)})_{ii}} \\ &= (x_0^{(k)})_{ii} - \frac{\left(1 - \frac{a_i}{(x_0^{(k)})_{ii}} - (x_0^{(k)})_{ii}\right)^2}{1 - \frac{a_i}{(x_0^{(k)})_{ii}} - 2\left(1 - \frac{a_i}{(x_0^{(k)})_{ii}}\right) + (x_0^{(k)})_{ii}}. \end{aligned}$$

Then, since $\frac{1}{2} \leq (x_0^{(k)})_{ii} \leq 1$ and $0 \leq a_i \leq \frac{1}{4}$, then we can find the maximum and minimum value of $(x_0^{(k+1)})_{ii}$. If we consider $(x_0^{(k+1)})_{ii}$ as a function of two variable $(x_0^{(k)})_{ii}$ and a_i , then by a simple calculation we have $(x_0^{(k+1)})_{ii} \in [\frac{1}{2}, 1]$, so $X_0^{(k+1)} \in [\frac{1}{2}I, I]$, and similarly we can show that $X_1^{(k+1)}$ and $X_2^{(k+1)} \in [\frac{1}{2}I, I]$. ■

4. Numerical example

Here, we present some examples and we will apply fixed point and Steffensen methods for them, and compare the number of iterations for those methods. Note that we have computed the numerical results by Maple software.

Example 4.1 Consider the matrix equation (2) with the normal matrix $A = \frac{1}{2}I_{n \times n}$. By [6] the exact solution is $X = \frac{1}{2}I$. We know that $\rho(A) = 0.5$, so the nonlinear matrix equation (2) has a solution. Comparison among fixed point method(FPM) and Steffensen method(SM) for $n = 2, 5, 10, 15, 20$ also iteration numbers(IN) and error(ERR) where

$ERR = \|X_k - X\|_2$ and X is the exact solution of the nonlinear matrix equation (2), with $A = \frac{1}{2}I_{n \times n}$ are showed in Table 1.

Numerical results for example 4.1

Method		n=2	n=5	n=10	n=15	n=20
FPM	ERR:	0.01	0.01	0.01	0.01	0.01
	IN:	51	51	51	51	51
	ERR:	0.001	0.001	0.001	0.001	0.001
	IN:	501	501	501 & 501	501	
	ERR:	0.0001	0.0001	0.0001	0.0001	0.0001
	IN:	5000	5000	5000	5000	5000
SM	ERR:	0.01	0.01	0.01	0.01	0.01
	IN:	5	5	5	5	5
	ERR:	0.001	0.001	0.001	0.001	0.001
	IN:	8	8	8	8	8
	ERR:	0.0001	0.0001	0.0001	0.0001	0.0001
	IN:	11	11	11	11	11

Example 4.2 Consider the matrix equation (2) with the normal matrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 0.5 \\ 0 & \dots & 0 & 0.5 & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & 0.5 & 0 & \dots & 0 \\ 0.5 & 0 & \dots & 0 & 0 \end{pmatrix}_{n \times n}$$

We know that $\rho(A) = 0.5$, so the nonlinear matrix equation (2) has a solution. By the ref. [6] the exact solution is $X = \frac{1}{2}I_n$. Comparison among fixed point method(FPM) and Steffensen method(SM) for $n = 2, 5, 10, 15, 20$ also iteration numbers(IN) and error(ERR) where $ERR = \|X_k - X\|_2$ and X is the exact solution of the nonlinear matrix equation (2), with A are showed in Table 2.

Numerical results for example 4.2

Method		n=2	n=5	n=10	n=15	n=20
FPM	ERR:	0.01	0.01	0.01	0.01	0.01
	IN:	51	51	51	51	51
	ERR:	0.001	0.001	0.001	0.001	0.001
	IN:	501	501	501 & 501	501	
	ERR:	0.0001	0.0001	0.0001	0.0001	0.0001
	IN:	5000	5000	5000	5000	5000
SM	ERR:	0.01	0.01	0.01	0.01	0.01
	IN:	5	5	5	5	5
	ERR:	0.001	0.001	0.001	0.001	0.001
	IN:	8	8	8	8	8
	ERR:	0.0001	0.0001	0.0001	0.0001	0.0001
	IN:	11	11	11	11	11

Example 4.3 Consider the matrix equation (2) with the normal matrix

$$A = \begin{pmatrix} 0.4 & 0 & 0 & -0.25 \\ 0 & 0.33 & -0.2 & 0 \\ 0 & -0.2 & -0.33 & 0 \\ -0.25 & 0 & 0 & -0.4 \end{pmatrix}.$$

We know that $\rho(A) = 0.4717$, so the nonlinear matrix equation (2) has a solution. By the ref. [6] the exact solution is,

$$X = \begin{pmatrix} 0.6658341 & 0 & 0 & 0 \\ 0 & 0.8144782 & 0 & 0 \\ 0 & 0 & 0.8144782 & 0 \\ 0 & 0 & 0 & 0.6658341 \end{pmatrix}.$$

The fixed point method needs 22 iterations to find the approximation solution with accuracy 0.000001 and the Steffensen method needs 4 iterations. Applying these iterative methods, yields

$$X = \begin{pmatrix} 0.665834 & 0 & 0 & 0 \\ 0 & 0.814478 & 0 & 0 \\ 0 & 0 & 0.814478 & 0 \\ 0 & 0 & 0 & 0.665834 \end{pmatrix}.$$

Example 4.4 Consider the matrix equation (2) with the normal matrix

$$A = \begin{pmatrix} 0.25 & 0 & 0.1 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0.1 & 0 & 0.1 & 0.2 \\ 0 & 0 & 0.2 & 0.25 \end{pmatrix}.$$

We know that $\rho(A) = 0.4108$, so the nonlinear matrix equation (2) has a solution. By the ref. [6] the exact solution is,

$$X = \begin{pmatrix} 0.9178145 & 0 & -0.0448002 & -0.0303964 \\ 0 & 0.9582574 & 0 & 0 \\ -0.0448002 & 0 & 0.9242223 & -0.0896007 \\ -0.0303964 & 0 & -0.0896007 & 0.87222026 \end{pmatrix}.$$

The fixed point method needs 11 iterations to find the approximation solution with accuracy 0.000001 and the Steffensen method needs 3 iterations. Applying these iterative methods, yields

$$X = \begin{pmatrix} 0.917814 & 0 & -0.044800 & -0.030396 \\ 0 & 0.958257 & 0 & 0 \\ -0.044800 & 0 & 0.924222 & -0.089600 \\ -0.030396 & 0 & -0.089600 & 0.8722202 \end{pmatrix}.$$

Example 4.5 Consider the matrix equation (2) with the normal matrix

$$A = \begin{pmatrix} -\frac{1}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{1}{7} \end{pmatrix}.$$

We know that $\rho(A) = 0.3194$, so the nonlinear matrix equation (2) has a solution. By the ref. [6] the exact solution is, $X = 0.8846458I$. The fixed point method needs 9 iterations to find the approximation solution with accuracy 0.000001 and the Steffensen method needs 3 iterations. Applying these iterative methods, yields $X = 0.884645I$.

Example 4.6 Consider the matrix equation (2) with the normal matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{3}{7} & 0 \\ 0 & 0 & -\frac{2}{5} & 0 & 0 \\ 0 & \frac{3}{7} & 0 & 0 & 0 \\ \frac{1}{5} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We know that $\rho(A) = 0.4286$, so the nonlinear matrix equation (2) has a solution. By the ref. [6] the exact solution is,

$$X = \begin{pmatrix} 0.9582575 & 0 & 0 & 0 & 0 \\ 0 & 0.7575451 & 0 & 0 & 0 \\ 0 & 0 & 0.8000005 & 0 & 0 \\ 0 & 0 & 0 & 0.7575451 & 0 \\ 0 & 0 & 0 & 0 & 0.9582575 \end{pmatrix}.$$

The fixed point method needs 17 iterations to find the approximation solution with accuracy 0.000001 and the Steffensen method needs 4 iterations. Applying these iterative methods, yields

$$X = \begin{pmatrix} 0.958257 & 0 & 0 & 0 & 0 \\ 0 & 0.757545 & 0 & 0 & 0 \\ 0 & 0 & 0.800000 & 0 & 0 \\ 0 & 0 & 0 & 0.757545 & 0 \\ 0 & 0 & 0 & 0 & 0.958257 \end{pmatrix}.$$

References

- [1] R. Bhatia, *Matrix analysis*, Springer, Berlin, 1997.
- [2] Burden, Richard. L, *Numerical analysis*, 6th, ed. 1997.
- [3] Fuzhen Zhang, *Matrix Theory Basic Result and Techniques*, Springer, 1999.
- [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 App. 194 (1993) 91-108.
- [5] S. M. El-Sayed, A. M. Al-Dbiban, *A new inversion free iteration for solving the equation $X + A^T X^{-1} A = Q$* , Linear Algebra App. 181 (2005) 148-156.
- [6] X. Zhan, J. Xie, *On the Matrix Equation $X + A^T X^{-1} A = I$* , Linear Algebra Appl. 247, (1996) 337-345.
- [7] D. V. Ouellette, *Schur complements and statistics*, Linear Algebra Appl. 36 (1981) 187-295.
- [8] W. N. Anderson, T. D. Morley, and G. E. Trapp, *Ladder networks, futed Points, and the geometric mean*, Circuits Systems Signal Process. 3 (1983) 259-268.
- [9] T. Ando, *Limit of cascade iteration of matrices*, Numer. Funct. Anal. Optim. 21 (1980) 579-589.
- [10] R. S. Bucy, *A priori bound for the Riccati equation*, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. III.

- [11] W. L. Green and E. Kamen, *Stabilization of linear Systems over a commutative normed algebra with applications to spatially distributed Parameter dependent Systems*, SIAM J. Control Optim. 23 (1985) 1-18.
- [12] W. Pusz and S. L. Woronowitz, *Funcitonal calculus for sequilinear forms and the purification map*, Rep. Math. Phys. 8 (1975) 159-170.
- [13] G. E. Trapp, *The Ricatii equation and the geometric mean*, Contemp. Math. 47 (1985) 437-445.
- [14] J. Zabezyk, *Remarks on the control of discrete time distributed paramter Systems*, SIAM J. Control 12 (1974) 721-735.
- [15] J. Zemanian, *Non-uniform semi-infinite grounded grids*, SIAM J. Appl. Math. 13 (1982) 770-788.
- [16] R. A. Horn and C. A. Johnson, *Matrix Analysis*, Cambridge, U.P. London, 1985.
- [17] R. Bhatia, C. Davis, *More matrix forms of the arithmetic-geometric mean inequality*, SIAM J. Matrix Anal. Appl. 14 (1993) 132-136.
- [18] J. C. Engwerda, A. C. M. Ran, and A. L. Rijkeboer, *Necessary and sufficient conditions for the existente of a positive definite solution of the matrix equation $X + A^*X^{-1}A = Q$* , Linear Algebra Appl. 186 (1993) 255-275 .

Archive of SID