Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 16, No. 2/2024

ON INEXACT NEWTON METHODS FOR SOLVING TWO NONLINEAR MATRIX EQUATIONS*

Karolina Nikolova† Vejdi I. Hasanov‡

Abstract

In this paper we consider inexact Newton methods for finding the largest positive definite solutions of two nonlinear matrix equations $X+A^*X^{-1}A=Q$ and $X-A^*X^{-1}A=Q$, respectively. Using Newton's method for considered equations requires solving a Stein's equation at each iteration. For solving the Stein's equation, we use Smith-type iterations instead of exact methods. Nonlocal convergence of the process is shown. Numerical experiments are included to illustrate the theory.

Keywords: nonlinear matrix equation, positive definite solution, inexact Newton method.

MSC: 65F10, 15A24.

DOI https://doi.org/10.56082/annalsarscimath.2024.2.311

1 Introduction

We consider iterative methods for solving the nonlinear matrix equations

$$X + A^*X^{-1}A = Q \tag{1}$$

*Accepted for publication on September 5, 2024

†karolinanikolova@abv.bg Konstantin Preslavsky University of Shumen, Faculty of Mathematics and Informatics, Shumen, Bulgaria

‡v.hasanov@shu.bg Konstantin Preslavsky University of Shumen, Faculty of Mathematics and Informatics, Shumen, Bulgaria

and

$$X - A^*X^{-1}A = Q \tag{2}$$

where Q is a $n \times n$ Hermitian positive definite matrix, A is a $n \times n$ matrix, and A^* denotes the conjugate transpose of A. The equations (1) and (2) can be reduced to the corresponding equations $X + A_1^*X^{-1}A_1 = I$ and $X - A_1^*X^{-1}A_1 = I$ with right-hand side I (I is identity matrix). The considered equations with right-hand side Q or I have been studied by many authors [1, 2, 3, 4, 5, 6, 7, 8]. Both equations find application in various fields, see references given in [1] and [5].

Eq. (1) was introduced by Anderson et al. [1]. Engwerda, Ran and Rijkeboer [2] have investigated the theoretical properties of Eq. (1) as necessary and sufficient conditions for the existence of a positive definite solution (PDS). They also have given conditions for a largest and smallest positive definite solutions (PDSs) and iterative methods (fixed-point iterations) for finding them. Moreover, it is proven the connection of (1) with a discrete algebraic Riccati equation [2, 3].

Eq. (2) and its Hermitian solutions were studied by Ferrante and Levy [5]. They have proved that the largest solution corresponds to its unique PDS. They also have studied its relationship with another type discrete algebraic Riccati equation and have given an iterative method (fixed-point iteration) for finding the unique PDS.

Besides the fixed-point iterations presented in [1, 2, 3] and [5] for equations (1) and (2), respectively, some authors have proposed other iterative methods for finding PDSs. Zhan [4] has proposed an inverse-free variant of the fixed-point iteration for the largest solution of Eq. (1). Guo and Lancaster [6] have proposed second inverse-free variant of the fixed-point iteration for Eq. (1) and Newton's method for both equations (1) and (2). Meini [7] has studied a cyclic reduction method for finding a largest and a smallest solutions of the equations (1) and (2), respectively.

Motivated by investigation of Gao and Bai [10], we consider inexact Newton methods for finding the largest solutions of the equations (1) and (2), respectively. It is known that the Newton's method requires solving a Stein's equation at each iteration. To solve it, instead of exact methods, we use 2-Smith iteration [9].

The paper is organized as follows: In Section 2, we present some known necessary and/or sufficient conditions for the existence of the PDS and some iterative methods for solving the considered equations. In Section 3 we consider inexact Newton methods and finally in Section 4 we will give numerical examples to show the behavior of the considered methods.

Throughout the paper $A > 0$ ($A \geq 0$) means that A is a Hermitian positive definite (semi-definite) matrix, $\rho(A)$ and $\|A\|$ are the spectral radius and the spectral norm of a matrix A, respectively. Moreover, for a $n \times n$ matrix A, we use $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$ We will say that X_S and X_L are a *smallest* and a *largest* PDSs of Eq. (1), respectively, when for every PDS X, we have $X_S \leq X \leq X_L$. Also, we will say that X_- and X_+ are a smallest and a largest Hermitian solutions of the Eq.(2), respectively, when for every Hermition solution X we have $X_- \leq X \leq X_+$.

2 Preliminaries

In this section, we present some known necessary and/or sufficient conditions for the existence of the PDS and some iterative methods for the largest PDS to the equations (1) and (2), respectively.

Engwerda et al. [2] have studied solvability of Eq. (1) in terms of properties of the corresponding rational matrix-valued function $\psi(\lambda) = Q + \lambda A + \lambda^{-1} A^*$. They have proved that: Under assumption $Q > 0$, Eq. (1) has a PDS if and only if ψ is regular (i.e. $\det \psi(\lambda) \neq 0$ for some λ) and $\psi(\lambda) \geq 0$ for all λ on the unit circle [2, Theorem 2.1]. Moreover:

Theorem 1. *[2, Theorem 3.4] Suppose $Q > 0$ and assume Eq. (1) has a PDS. Then this equation has a largest and a smallest solution X_L and X_S, respectively. Moreover, X_L is the unique solution for which $X + \lambda A$ is invertible for $|\lambda| < 1$, while X_S is the unique solution for which $X + \lambda A^*$ is invertible for $|\lambda| > 1$.*

Therefore, the largest PDS X_L of Eq. (1) is the unique solution for which $\rho(X_L^{-1} A) \leq 1$.

Let us denote by $\omega(A)$ the numerical radius of a matrix A, i.e.,

$$\omega(A) = \max_{\|x\|=1} |x^* A x|.$$

Theorem 2. *[2, Theorem 5.2] Suppose A is nonsingular. Then Eq. (1) has a PDS X if and only if $\omega(Q^{-\frac{1}{2}} A Q^{-\frac{1}{2}}) \leq \frac{1}{2}$.*

Remark 1. *Let $Q = L^T L$ be Cholesky decomposition of $Q > 0$. Then $Q^{-\frac{1}{2}} A Q^{-\frac{1}{2}}$ can be replaced with $L^{-T} A Q^{-1}$ in Theorem 2.*

Ferrante and Levy [5] have investigated Eq. (2). Following are some of their results:

Theorem 3. *[5, Proposition 4.1] The set of solutions of Eq. (2) admits a largest element, which is the unique PDS X_+.*

Theorem 4. *[5, Corollary 4.1] The set of solutions of Eq. (2) admits a smallest element X_-, which for A nonsingular is its unique negative definite solution.*

Now we will present some known iterative methods for founding the largest solutions X_L and X_+ of Eq. (1) and Eq. (2), respectively, which we will compare with our method.

Firstly, we present the iterative methods for Eq. (1).

Algorithm 1 (Basic fixed-point iteration for Eq. (1))**.** *Take $X_0 = Q$. For $k = 0, 1, \ldots,$ compute*

$$X_{k+1} = Q - A^* X_k^{-1} A,$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, for a fixed error bound $tol > 0$.

Then $X_L \approx X_{k+1}$.

For Algorithm 1 (BFPI-(1)) we have:

- $X_0 \geq X_1 \geq X_2 \geq \cdots$, and $\lim_{k\to\infty} X_k = X_L$ [2];
- $\lim_{k\to\infty} sup \sqrt[k]{\|X_k - X_L\|} \leq (\rho(X_L^{-1}A))^2$ [6, Theorem 2.3].

We mentioned earlier that $\rho(X_L^{-1}A) \leq 1$. From above, the convergence of the BFPI-(1) is R-linear, whenever $\rho(X_L^{-1}A) < 1$, and typically sublinear, when $\rho(X_L^{-1}A) = 1$.

Next algorithm, representing Newton's method was studied by Guo and Lancaster [6].

Algorithm 2 (Newton's method for Eq. (1))**.** *Take $X_0 = Q$. For k=0,1,..., compute $L_k = X_k^{-1}A$, and solve*

$$X_{k+1} - L_k^* X_{k+1} L_k = Q - 2L_k^* A, \tag{3}$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, for a fixed error bound $tol > 0$.

Then $X_L \approx X_{k+1}$.

Algorithm 2 (NM-(1)) has the following properties [6, Theorem 5.3]:

- $\rho(L_k) < 1$, for $k = 0, 1, \ldots$;
- $X_0 \geq X_1 \geq X_2 \geq \cdots$, and $\lim_{k\to\infty} X_k = X_L$;

- The convergence is quadratic if $\rho(X_L^{-1}A) < 1$. If $\rho(X_L^{-1}A) = 1$ and all eigenvalues of $X_L^{-1}A$ on the unit circle are semisimple, then the convergence is either quadratic or linear with rate $1/2$.

The third method we will present is the Cyclic reduction method, considered by Meini [7].

Algorithm 3 (Cyclic reduction for Eq. (1))**.** *Set* $X_0 = Y_0 = Q_0 = Q$, $A_0 = A$. *For* $k = 0, 1, \ldots$, *compute*

$$\begin{aligned} A_{k+1} &= A_k Q_k^{-1} A_k, \\ Q_{k+1} &= Q_k - A_k Q_k^{-1} A_k^* - A_k^* Q_k^{-1} A_k, \\ X_{k+1} &= X_k - A_k^* Q_k^{-1} A_k, \\ Y_{k+1} &= Y_k - A_k Q_k^{-1} A_k^*, \end{aligned}$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, *for a fixed error bound* $tol > 0$.

Then $X_L \approx X_{k+1}$ *and* $X_S \approx Y_{k+1}$.

The Meini's cyclic reduction method has the following properties [7]:

- $Q_0 \geq Q_1 \geq Q_2 \geq \cdots > 0$;
- $X_0 \geq X_1 \geq X_2 \geq \cdots$, and $\lim_{k\to\infty} X_k = X_L$;
- $Y_0 \geq Y_1 \geq Y_2 \geq \cdots$, $\lim_{k\to\infty} Y_k = Y_L$, and $X_S = Q - Y_L$;
- The convergence is quadratic if $\rho(X_L^{-1}A) < 1$. If $\rho(X_L^{-1}A) = 1$ and all eigenvalues of $X_L^{-1}A$ on the unit circle are semisimple, then the convergence is at least linear with rate $\frac{1}{2}$ [8].

Now we will present analogous methods to the above for Eq. (2).

Algorithm 4 (Basic fixed-point iteration for Eq. (2))**.** *Take* $X_0 = Q$. *For* $k = 0, 1, \ldots$, *compute*

$$X_{k+1} = Q + A^* X_k^{-1} A,$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, *for a fixed error bound* $tol > 0$.

Then $X_+ \approx X_{k+1}$.

For Algorithm 4 (BFPI-(2)), we have:

- $X_{2k-2} \leq X_{2k} \leq X_{2s+1} \leq X_{2s-1}$, for all $k, s \geq 1$, and $\lim_{k\to\infty} X_k = X_+$ [5, Lemma 5.1];

- $\lim_{k\to\infty} sup \sqrt[k]{\|X_k - X_+\|} \leq (\rho(X_+^{-1}A))^2 < 1$ [6, Theorem 2.6].

Hence, the convergence of the BFPI-(2) is R-linear.

Algorithm 5 (Newton's method for Eq. (2))**.** *For given X_0 sufficiently close to X_+ and k=0,1,..., compute $L_k = X_k^{-1}A$, and solve*

$$X_{k+1} + L_k^* X_{k+1} L_k = Q + 2L_k^* A, \tag{4}$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, for a fixed error bound $tol > 0$.

Then $X_+ \approx X_{k+1}$.

Newton's method for Eq (2) guarantees quadratic convergence if the initial guess X_0 must be chosen close to the solution X_+. It can be used as an efficient correction method [6].

Algorithm 6 (Cyclic reduction for Eq. (2))**.** *Set*

$$\begin{aligned} &A_1 = AQ^{-1}A, \quad Q_1 = Q + AQ^{-1}A^* + A^*Q^{-1}A, \\ &X_1 = Q + A^*Q^{-1}A, \quad Y_1 = Q + AQ^{-1}A^*. \end{aligned}$$

For $k = 1, 2, \ldots,$ compute

$$\begin{aligned} A_{k+1} &= A_k Q_k^{-1} A_k, \\ Q_{k+1} &= Q_k - A_k Q_k^{-1} A_k^* - A_k^* Q_k^{-1} A_k, \\ X_{k+1} &= X_k - A_k^* Q_k^{-1} A_k, \\ Y_{k+1} &= Y_k - A_k Q_k^{-1} A_k^*, \end{aligned}$$

until $\|X_{k+1} - X_k\|_\infty \leq tol$, for a fixed error bound $tol > 0$.

Then $X_+ \approx X_{k+1}$ and $X_- \approx Y_{k+1}$.

3 Inexact Newton methods

3.1 Inexact Newton methods for Eq. (1)

We remind that the Newton's method for Eq. (1) requires solving Eq. (3), which we rewrite in the following way:

$$(X_{k+1} - X_k) - L_k^*(X_{k+1} - X_k)L_k = Q - L_k^* A - X_k. \tag{5}$$

We note that Eq. (5) (i.e. (3)) is of the type of the well-known Stein's equation:

$$Y - C^* Y C = D. \tag{6}$$

Lemma 1. *[11] Let C, D be square matrices.*

(a) If $\rho(C) < 1$, then Eq. (6) has a unique solution Y, and $Y \geq 0$ $(Y > 0)$, when $D \geq 0$ $(D > 0)$.

(b) If there is some $Y > 0$ such that $Y - C^*YC$ is positive definite (semi-definite), then $\rho(C) < 1$ $(\rho(C) \leq 1)$.

Remark 2. *In case of $\rho(C) < 1$ the unique solution Y of Eq. (6) has representation*

$$Y = \sum_{j=0}^{\infty} (C^*)^j D C^j.$$

Remark 3. *Under condition of Lemma 1, (a), we have $Y \leq 0$ $(Y < 0)$, when $D \leq 0$ $(D < 0)$.*

The Stein's equation (6) can be solved with an exact method or iteratively.

Further in a considered method, similar to Gao and Bai [10], we will use 2-Smith iteration [9]: Set $Y_0 = D$, $C_0 = C$. For $p = 0, 1, \ldots,$ compute

$$Y_{p+1} = Y_p + C_p^* Y_p C_p, \quad C_{p+1} = C_p^2. \tag{7}$$

We have that

$$Y_p = \sum_{j=0}^{2^p - 1} (C^*)^j D C^j, \quad p = 0, 1, \ldots.$$

Hence, in case of $\rho(C) < 1$ the sequence $\{Y_p\}$ quadratically converges to the unique solution Y of Eq. (6).

Now we propose two variants of the Inexact Newton method for Eq. (1). Firstly,

Algorithm 7 (Inexact Newton method for Eq. (1))**.** *Take $X_0 = Q$ and m.*

$$
\begin{aligned}
& For\ k = 0, 1, \ldots,\ compute \\
& L_k = X_k^{-1} A, \quad D_k = Q - L_k^* A - X_k, \\
& \qquad take\ C_{k,0} = L_k, \quad Y_{k,1} = D_k + C_{k,0}^* D_k C_{k,0}, \\
& \qquad for\ p = 1, 2, \ldots, m,\ compute \\
& \qquad C_{k,p} = C_{k,p-1}^2, \quad Y_{k,p+1} = Y_{k,p} + C_{k,p}^* Y_{k,p} C_{k,p}, \\
& X_{k+1} = X_k + Y_{k,m+1}, \\
& until\ \|Y_{k,m+1}\|_\infty \leq tol,\ for\ a\ fixed\ error\ bound\ tol > 0.
\end{aligned}
$$

Then $X_L \approx X_{k+1}$.

In the each step k in the Algorithm 7 we approximate Y_k of the equation

$$Y_k - L_k^* Y_k L_k = D_k \tag{8}$$

by $Y_{k,m+1}$ (see Eq. (5)), using the 2-Smith iteration.

In the second variant of the Inexact Newton method (INM-(1)), the fixed number m of the 2-Smith iteration, we replace with k − the iteration step.

Algorithm 8 (Inexact Newton method for Eq. (1))**.** *Take $X_0 = Q$.*

$$\begin{aligned}
&For\ k = 0, 1, \ldots,\ compute\\
&L_k = X_k^{-1} A, \quad D_k = Q - L_k^* A - X_k,\\
&\qquad take\ C_{k,0} = L_k, \quad Y_{k,1} = D_k + C_{k,0}^* D_k C_{k,0},\\
&\qquad for\ p = 1, 2, \ldots, k,\ compute\\
&\qquad C_{k,p} = C_{k,p-1}^2, \quad Y_{k,p+1} = Y_{k,p} + C_{k,p}^* Y_{k,p} C_{k,p},\\
&X_{k+1} = X_k + Y_{k,k+1},\\
&until\ \|Y_{k,k+1}\|_\infty \le tol,\ for\ a\ fixed\ error\ bound\ tol > 0.
\end{aligned}$$

Then $X_L \approx X_{k+1}$.

Theorem 5. *If Eq. (3) has a positive definite solution, then Algorithm 7 determines a sequence of Hermitian matrices $\{X_k\}$ for which $\rho(L_k) < 1$, $D_k \le 0$, $X_k \ge X_{k+1}$ for $k = 0, 1, \ldots$, and $\lim_{k\to\infty} X_k = X_L$, where X_L is the largest solution of Eq. (3).*

Proof. Let X_+ be any PDS of Eq. (3), i.e. $Q - A^* X_+^{-1} A - X_+ = 0$.

We will prove the theorem by induction. Together with the sequence $\{X_k\}_{k=0}^\infty$, we consider the sequences $\{X'_k\}_{k=1}^\infty$ and $\{Y_k\}_{k=0}^\infty$, where Y_k is a unique solution of Eq. (8) and $X'_{k+1} = X_k + Y_k$.

After simple calculations for X'_{k+1}, we have

$$X'_{k+1} - L_k^* X'_{k+1} L_k = Q - 2 L_k^* A \tag{9}$$

and

$$X'_{k+1} - X_+ - L_k^* (X'_{k+1} - X_+) L_k = (A - X_+ L_k)^* X_+^{-1} (A - X_+ L_k). \tag{10}$$

For $k = 0$, we have $X_0 = Q \ge X_+$ and

$$0 < X_+ = Q - A^* X_+^{-1} A \le Q - A^* Q^{-1} A = Q - (Q^{-1} A)^* Q Q^{-1} A.$$

Thus, $\rho(L_0) < 1$ by Lemma 1.

By $D_0 = Q - A^*Q^{-1}A - Q = -A^*Q^{-1}A \leq 0$, Remark 2, and Remark 3, we have $Y_0 \leq Y_{0,m+1} \leq 0$ and $X_1' \leq X_1 \leq X_0$.

Since $\rho(L_0) < 1$, from (10) in case of $k = 0$ and Lemma 1, it follows $X_1' \geq X_+$. Hence $X_1 \geq X_+$.

Assume that the statement of the theorem is true for $k = q$, i.e.

$$\rho(L_q) < 1, \quad D_q \leq 0, \quad X_q \geq X_{q+1},$$

and we suppose $X_q \geq X_+$, also.

Now we will prove that it is true for $k = q + 1$.

From $\rho(L_q) < 1$, $D_q \leq 0$, and (10) in case of $k = q + 1$ we have

$$Y_q \leq Y_{q,m+1} \leq 0 \text{ and } X_+ \leq X'_{q+1} \leq X_{q+1} \leq X_q.$$

By using equalities

$$Y_q - Y_{q,m+1} - L_q^*(Y_q - Y_{q,m+1})L_q = (L_q^*)^{2^{m+1}} D_q L_q^{2^{m+1}},$$

$$(L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1}) = L_q^* X_{q+1} L_q - 2L_q^* A + L_{q+1}^* A,$$

and (9) in case of $k = q$, we have

$$\begin{aligned}
D_{q+1} &= Q - L_{q+1}^* A - X_{q+1} \\
&= X'_{q+1} - L_q^* X'_{q+1} L_q + 2L_q^* A - L_{q+1}^* A - X_{q+1} \\
&= Y_q - Y_{q,m+1} - L_q^*(Y_q - Y_{q,m+1})L_q - (L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1}) \\
&= (L_q^*)^{2^{m+1}} D_q L_q^{2^{m+1}} - (L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1}) \leq 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
X_{q+1} - L_{q+1}^* X_{q+1} L_{q+1} &= Q - 2L_{q+1}^* A - D_{k+1} \\
&= Q - 2L_{q+1}^* A - (L_q^*)^{2^{m+1}} D_q L_q^{2^{m+1}} \\
&\quad + (L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1})
\end{aligned}$$

and

$$\begin{aligned}
&X_{q+1} - X_+ - L_{q+1}^*(X_{q+1} - X_+)L_{q+1} \\
&\qquad = (A - X_+ L_{q+1})^* X_+^{-1}(A - X_+ L_{q+1}) - (L_q^*)^{2^{m+1}} D_q L_q^{2^{m+1}} \\
&\qquad\quad + (L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1}).
\end{aligned} \tag{11}$$

Let us consider the operators $\mathcal{S}_k : \mathcal{H}^n \to \mathcal{H}^n$, $k = 0, 1, \ldots$, related

$$\mathcal{S}_k(H) = L_k^* H L_k, \quad H \in \mathcal{H}^n, \tag{12}$$

where $\mathcal{H}^n$ is the space of $n \times n$ Hermitian matrices, endowed with the scalar product $\langle X, Y\rangle = \operatorname{trace} XY$. We note that the spectrum of $\mathcal{S}_k$ is equal to the spectrum of $L_k^T \otimes L_k^*$ ($\otimes$ means the Kronecker product) [12]. So $\rho(\mathcal{S}_k) < 1$ if and only if $\rho(L_k) < 1$. Hence $\rho(\mathcal{S}_q) < 1$.

Assume that $\rho(\mathcal{S}_{q+1}) \geq 1$. Since $\mathcal{S}_{q+1}$ is positive operator, defined on $\mathcal{H}^n$, by Krein-Rutman's theorem [13] it follows that there is $V \geq 0$, $V \neq 0$ with $\mathcal{S}_{q+1}^*(V) = \lambda V$ for some $\lambda \geq 1$. Let W be the right-hand side of (11), then we get

$$\begin{aligned} 0 \leq \langle V, W\rangle &= \langle V, X_{q+1} - X_+\rangle - \langle V, \mathcal{S}_{q+1}(X_{q+1} - X_+)\rangle \\ &= (1-\lambda)\langle V, X_{q+1} - X_+\rangle \leq 0. \end{aligned}$$

Hence $\langle V, W\rangle = 0$. Since $D_q \leq 0$, $X_+^{-1} > 0$, $X_{q+1} > 0$ in (11), we conclude that $\langle V, (L_q - L_{q+1})^* X_{q+1}(L_q - L_{q+1})\rangle = 0$ and $(L_q - L_{q+1})V = 0$, i.e., $L_q V = L_{q+1} V$. Thus, $\mathcal{S}_q^*(V) = \mathcal{S}_{q+1}^*(V) = \lambda V$, which contradicts $\rho(\mathcal{S}_q) < 1$. So, $\rho(\mathcal{S}_{q+1}) < 1$ and $\rho(L_{q+1}) < 1$.

Therefore, $\rho(L_k) < 1$, $D_k \leq 0$, and $X_k \geq X_{k+1} \geq X_+$ for $k = 0, 1, \ldots q$, where X_+ is arbitrary PDS of Eq. (1). Thus, $\lim_{k\to\infty} = X_L$. □

Theorem 6. *If Eq. (3) has a positive definite solution, then Algorithm 8 determines a sequence of Hermitian matrices* $\{X_k\}$ *for which* $\rho(L_k) < 1$, $D_k \leq 0$, $X_k \geq X_{k+1}$ *for* $k = 0, 1, \ldots,$ *and* $\lim_{k\to\infty} X_k = X_L$, *where* X_L *is the largest solution of Eq. (3).*

Proof. The proof is analogous to Theorem 5. □

3.2 Inexact Newton methods for Eq. (2)

For Eq. (2) the Newton's method requires solving Eq. (4), which we rewrite in the following way:

$$(X_{k+1} - X_k) + L_k^*(X_{k+1} - X_k)L_k = Q + L_k^* A - X_k. \tag{13}$$

Both of the equations (4) and (13) are of the following type equation:

$$Y + C^* Y C = D. \tag{14}$$

A sufficient condition for the existence of a unique solution $\bar{Y}$ of Eq. (14) is $\tilde{D} - C^* \tilde{D} C > 0$ for a $\tilde{D} > 0$ (see [12, Proposition 3.1]). Thus, by Lemma 1 it follows $\rho(C) < 1$. Moreover, in this case the unique solution is given by

$$\bar{Y} = \sum_{j=0}^{\infty} (-1)^j (C^*)^j D C^j. \tag{15}$$

From $D > 0$ does not follow $\bar{Y} > 0$, but if $D - C^*DC > 0$, then $\bar{Y} > 0$ (see [12, Theorem 3.3]). The unique solution $\bar{Y}$ of Eq. (14) can be computed iteratively by 2-Smith iteration (7) with $Y_1 = D - C^*DC$ and $C_1 = C^2$.

Now we propose analogous of Algorithm 7 and Algorithm 8 variants of the Inexact Newton method for Eq. (2).

Algorithm 9 (Inexact Newton method for Eq. (2))**.** *Take X_0 sufficiently close to X_+ and m.*

$$
\begin{aligned}
&For\ k = 0, 1, \dots,\ compute\\
&L_k = X_k^{-1}A, \quad D_k = Q + L_k^* A - X_k,\\
&\qquad take\ C_{k,0} = L_k, \quad Y_{k,1} = D_k - C_{k,0}^* D_k C_{k,0}\\
&\qquad for\ p = 1, 2, \dots, m,\ compute\\
&\qquad C_{k,p} = C_{k,p-1}^2, \quad Y_{k,p+1} = Y_{k,p} + C_{k,p}^* Y_{k,p} C_{k,p},\\
&X_{k+1} = X_k + Y_{k,m+1},\\
&until\ \|Y_{k,m+1}\|_\infty \le tol,\ for\ a\ fixed\ error\ bound\ tol > 0.
\end{aligned}
$$

Then $X_+ \approx X_{k+1}$.

Algorithm 10 (Inexact Newton method for Eq. (2))**.** *Take X_0 sufficiently close to X_+.*

$$
\begin{aligned}
&For\ k = 0, 1, \dots,\ compute\\
&L_k = X_k^{-1}A, \quad D_k = Q + L_k^* A - X_k,\\
&\qquad take\ C_{k,0} = L_k, \quad Y_{k,1} = D_k - C_{k,0}^* D_k C_{k,0}\\
&\qquad for\ p = 1, 2, \dots, k,\ compute\\
&\qquad C_{k,p} = C_{k,p-1}^2, \quad Y_{k,p+1} = Y_{k,p} + C_{k,p}^* Y_{k,p} C_{k,p},\\
&X_{k+1} = X_k + Y_{k,k+1},\\
&until\ \|Y_{k,k+1}\|_\infty \le tol,\ for\ a\ fixed\ error\ bound\ tol > 0.
\end{aligned}
$$

Then $X_+ \approx X_{k+1}$.

In the two algorithms above, X_0 sufficiently close to X_+ is needed, as in the original Newton's method (Algorithm 5).

We present these algorithms by numerical examples in the next section and its convergence behavior we investigate in a future work.

Now we will propose an algorithm for Eq. (2) by using its relation to an equation of the type of Eq. (1).

Let X_+ be a unique PDS of Eq. (2). Then, by Woodbury formula, we have

$$\begin{aligned} X_+ &= Q + A^*X_+^{-1}A = Q + A^*(Q + A^*X_+^{-1}A)^{-1}A \\ &= Q + A^*[Q^{-1} - Q^{-1}A^*(X_+ + AQ^{-1}A^*)^{-1}AQ^{-1}]A \\ &= Q + A^*Q^{-1}A - A^*Q^{-1}A^*(X_+ + AQ^{-1}A^*)^{-1}AQ^{-1}A \end{aligned}$$

and

$$Z_L + B^*Z_L^{-1}B = P,$$

where $B = AQ^{-1}A$, $P = Q + A^*Q^{-1}A + AQ^{-1}A^*$, and $Z_L = X_+ + AQ^{-1}A^*$ is the largest solution of equation $Z + B^*Z^{-1}B = P$. Thus, following Algorithm 8, we get:

Algorithm 11 (Inexact Newton method for Eq. (2)). *Take* $Z_0 = Q + A^*Q^{-1}A + AQ^{-1}A^*$.

$$\begin{aligned} &For\ k = 0, 1, \ldots,\ compute \\ &L_k = Z_k^{-1}A, \quad D_k = Z_0 - L_k^*A - Z_k, \\ &\qquad take\ C_{k,0} = L_k, \quad Y_{k,1} = D_k + C_{k,0}^*D_kC_{k,0}, \\ &\qquad for\ p = 1, 2, \ldots, k,\ compute \\ &\qquad C_{k,p} = C_{k,p-1}^2, \quad Y_{k,p+1} = Y_{k,p} + C_{k,p}^*Y_{k,p}C_{k,p}, \\ &Z_{k+1} = Z_k + Y_{k,k+1}, \\ &until\ \|Y_{k,k+1}\|_\infty \le tol,\ for\ a\ fixed\ error\ bound\ tol > 0. \end{aligned}$$

Then $X_+ \approx Z_{k+1} - AQ^{-1}A^*$.

4 Numerical examples

In this section, we conduct numerical experiments to approximate the largest PDS of Eq. (1) and the unique PDS of Eq. (2), respectively. We compare the proposed Inexact Newton methods with Basic fixed-point iteration, Newton's method and Cyclic reduction for the both equations, respectively. As practical stopping criterion for each algorithm, we use $\|X_k - X_{k-1}\|_\infty \le tol$ for given $tol > 0$ ($\|Z_k - Z_{k-1}\|_\infty \le tol$ for Algorithm 11). For the considered examples and algorithms, we report k – the number of iterations, $\|X_k - X_{k-1}\|_\infty$, and $res_+(X) = \|Q - A^*X^{-1}A - X\|_\infty$ or $res_-(X) = \|Q + A^*X^{-1}A - X\|_\infty$.

Example 1. *[6, Example 7.1] We consider Eq. (1) with matrix coefficients*

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}, \qquad Q = \begin{pmatrix} 6.0 & 5.0 \\ 5.0 & 8.6 \end{pmatrix}.$$

For the matrices in Example 1, we have $\rho(Q^{-\frac{1}{2}}AQ^{-\frac{1}{2}}) \approx 0.434$. In Table 1 we report the results of experiments for Example 1.

Table 1: Numerical results of Example 1 with $tol = 10^{-8}$.

Algorithm	k	$\|X_k - X_{k-1}\|_\infty$	$res_+(X_k)$
Algorithm 1 (BFPI-(1))	27	$8.5492e-09$	$3.2977e-09$
Algorithm 3 (CR-(1))	6	$3.5822e-11$	$2.6645e-15$
Algorithm 2 (NM-(1))	6	$5.1056e-11$	$3.1508e-11$
Algorithm 8 (INM-(1))	6	$5.3001e-11$	$1.3323e-15$
Algorithm 7 (INM-(1)) (m=10)	6	$5.1068e-11$	$8.8818e-16$
Algorithm 7 (INM-(1)) (m=4)	6	$5.1063e-11$	$4.4409e-16$

Example 2. *[6, Example 7.2] We consider Eq. (1) with $Q = I$ and*

$$A = \begin{pmatrix} 0.2 & 0.2 & 0.1 \\ 0.2 & 0.15 & 0.15 \\ 0.1 & 0.15 & 0.25 \end{pmatrix}.$$

We note that in Example 2 the matrix A is normal and $\rho(A) = \frac{1}{2}$. The results of the experiment are presented in Table 2.

Table 2: Numerical results of Example 2 with $tol = 10^{-8}$

Algorithm	k	$\|X_k - X_{k-1}\|_\infty$	$res_+(X_k)$
Algorithm 1 (BFPI-(1))	7071	$9.9988e-09$	$9.9960e-09$
Algorithm 3 (CR-(1))	26	$7.5853e-09$	$8.3267e-17$
Algorithm 2 (NM-(1))	25	$8.1568e-09$	$1.6653e-16$
Algorithm 8 (INM-(1))	25	$9.6333e-09$	$2.2204e-16$
Algorithm 7 (INM-(1)) (m=10)	167	$9.9060e-09$	$4.8064e-12$
Algorithm 7 (INM-(1)) (m=4)	200	$4.0684e-07$	$1.2647e-08$

For Example 2 we notice that, with Algorithm 7 (INM-(1)), the number of iterations k grows when the number of internal iterations decreases. The iterations of Algorithm 8 coincides with those of NM-(1) and close to those of CR-(1). The Algorithm 7 in case of $m = 4$ terminated at 200 iterations.

Example 3. *[6, Example 7.3] We consider Eq. (1) with*

$$A = \begin{pmatrix} 0.37 & 0.13 & 0.12 \\ -0.30 & 0.34 & 0.12 \\ 0.11 & -0.17 & 0.29 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1.20 & -0.30 & 0.10 \\ -0.30 & 2.10 & 0.20 \\ 0.10 & 0.20 & 0.65 \end{pmatrix}.$$

For the matrices in Example 3, we have $\rho(Q^{-\frac{1}{2}}AQ^{-\frac{1}{2}}) \approx 0.487$. The results of the experiment are presented in Table 3.

Table 3: Numerical results of Example 3 with $tol = 10^{-12}$

Algorithm	k	$\|X_k - X_{k-1}\|_\infty$	$res_+(X_k)$
Algorithm 1 (BFPI-(1))	332	$9.4835e-13$	$8.8862e-13$
Algorithm 3 (CR-(1))	10	$1.2216e-16$	$3.3307e-16$
Algorithm 2 (NM-(1))	9	$1.3942e-15$	$1.1796e-16$
Algorithm 8 (INM-(1))	9	$1.5293e-15$	$1.1796e-16$
Algorithm 7 (INM-(1)) (m=10)	9	$1.0292e-15$	$1.1102e-16$
Algorithm 7 (INM-(1)) (m=4)	16	$9.3708e-13$	$8.2226e-15$

The following examples are for Eq. (2).

Example 4. *[6, Example 7.4] We consider Eq. (2) with matrix coefficients*

$$A = \begin{pmatrix} 50 & 20 \\ 10 & 60 \end{pmatrix}, \qquad Q = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}.$$

For the matrices in Example 4, we have $\rho(Q^{-1}A) \approx 27$. Hence, Algorithm 9 and Algorithm 10 with $X_0 = Q$ are not applicable. The results of the experiment are presented in Table 4. Indeed, Algorithm 9 and Algorithm 10 with $X_0 = Q$ are not convergent and terminated after 200 iterations.

Table 4: Numerical results of Example 4 with $tol = 10^{-10}$

Algorithm	k	$\|X_k - X_{k-1}\|_\infty$	$res_-(X_k)$
Algorithm 4 (BFPI-(2))	501	$9.4595e-11$	$8.7184e-11$
Algorithm 6 (CR-(2))	9	$2.4412e-11$	$2.3004e-12$
Algorithm 5 (NM-(2))	10	$1.0658e-14$	$1.4211e-14$
Algorithm 10 (INM-(2))	200	NaN	NaN
Algorithm 9 (INM-(2)) (m=10)	200	NaN	NaN
Algorithm 11 (INM-(2))	9	$4.2235e-13$	$2.0961e-12$

After 63 iterations by Algorithm 4 for Example 4, we have $\|(X_{62}^{-1}A\| + \|(X_{63}^{-1}A\| < 2$ and $\|(X_{63}^{-1}A\| \approx 0.95$. Then, using X_{63} as an initial value in Algorithm 5, Algorithm 9 and Algorithm 10, we obtain the results in Table 5.

Table 5: Numerical results of Example 4 after 63 iterations with BFPI-(2)

Algorithm	k	$\|X_{63+k} - X_{62+k}\|_\infty$	$res(X_{63+k})$
Algorithm 5 (NM-(2))	4	$2.4524e-11$	$1.7764e-14$
Algorithm 10 (INM-(2))	8	$1.7053e-13$	$1.0658e-14$
Algorithm 9 (INM-(2)) (m=10)	4	$2.4523e-11$	$1.7764e-14$
Algorithm 9 (INM-(2)) (m=4)	14	$1.5437e-11$	$5.4747e-12$

Example 5. *[14, Example 5] We consider Eq. (2) with $Q = I$ and*

$$A = \begin{pmatrix} -3.47 & 3.47 \\ -2.89 & -3.47 \end{pmatrix}.$$

For the matrix A in Example 5, we have $\rho(A) \approx 4.7$. Therefore, Algorithm 9 and Algorithm 10 with $X_0 = I$ are again not applicable. The results of the experiment are presented in Table 6.

Table 6: Numerical results of Example 5 with $tol = 10^{-10}$

Algorithm	k	$\|X_k - X_{k-1}\|_\infty$	$res_-(X_k)$
Algorithm 4 (BFPI-(2))	122	$9.4068e-11$	$7.7817e-11$
Algorithm 6 (CR-(2))	7	$1.4991e-11$	$1.1435e-14$
Algorithm 5 (NM-(2))	8	$1.1102e-16$	$1.2768e-15$
Algorithm 11 (INM-(2))	6	$3.4971e-11$	$6.6613e-15$

Algorithm 9 and Algorithm 10 we use with initial value X_7, which is obtained by Algorithm 4 (BFPI-(2)). The results are presented in Table 7.

Table 7: Numerical results of Example 5 after 7 iterations with BFPI-(2)

Algorithm	k	$\|X_{7+k} - X_{6+k}\|_\infty$	$res_-(X_{7+k})$
Algorithm 10 (INM-(2))	6	$7.3764e-12$	$1.2768e-15$
Algorithm 9 (INM-(2)) (m=10)	5	$9.6229e-13$	$1.3323e-15$
Algorithm 9 (INM-(2)) (m=4)	6	$3.2048e-11$	$6.4670e-14$

Acknowledgement. This work was partially supported by the Konstantin Preslavsky University of Shumen under project No. RD-08-144/02.02.2024.

References

[1] W.N. Anderson, T.D. Morley, and G.E. Trapp, Positive Solution to $X = A - BX^{-1}B^*$, *Linear Algebra Appl.* 134 (1990), 53-62.

[2] J.C. Engwerda, A.C.M. Ran and 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.

[3] J.C. Engwerda, On the Existence of a Positive Definite Solution of the Matrix Equation $X + A^TX^{-1}A = I$, *Linear Algebra Appl.* 194 (1993), 91-108.

[4] X.Z. Zhan Computing the extreme positive definite solutions of a matrix equation, *SIAM J. Sci. Comput.* 17 (1996), 632-645.

[5] A. Ferrante, B.C. Levy, Hermitian solutions of the equation $X = Q + NX^{-1}N^*$, *Linear Algebra Appl.* 247 (1996), 359-373.

[6] C.-H. Guo, P. Lancaster, Iterative Solution of Two Matrix Equations, *Math. Comput.* 68 (1999), 1589-1603.

[7] B. Meini, Efficient computation of the extreme solutions of $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$, *Math. Comput.* 71 (2001), 1189-1204.

[8] C.-H. Guo, Convergence rate of an iterative method for a nonlinear matrix equation, *SIAM J. Matrix Anal. Appl.* 23 (2001), 295-302.

[9] B. Zhou, J. Lam, G.-R. Duan, On Smith-type iterative algorithms for the Stein matrix equation, *Appl. Math. Letters* 22 (2009), 1038-1044.

[10] Y.-H. Gao, Z.-Z. Bai, On inexact Newton methods based on doubling iteration scheme for non-symmetric algebraic Riccati equations, *Numer. Linear Algebra Appl.* 18 (2011), 325-341.

[11] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, 2nd ed. San Diego (CA), Academic Press, 1985.

[12] A.C.M. Ran, M.C.B. Reurings, The Symmetric Matrix Equation, *The Electronic Journal of Linear Algebra* 9 (2002), 93-107.

[13] M.G. Krein, M.A. Rutman, Linear operators leaving invariant a cone in a Banach space, *Uspehi Matem. Nauk* 3(1) (1948), 3-95. (in Russian)

[14] I.G. Ivanov, V.I. Hasanov, F. Uhlig, Improved Methods and Starting Values to Solve the Matrix Equations $X \pm A^*X^{-1}A = I$ Iteratively, *Math. Comput.* 74 (2005), 263-278.