

Convergence of Relaxed-relaxed Newton Method method for Order-Convex Matrix Equations order-convex matrix equations

Copyright-Holder: SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional

Copyright-Year: 2019

Sang-hyup Seo ^{1*}, Jong-Hyeon Seo ²

¹Department of Mathematics, Pusan National University, Busan, 46241, Republic of Korea

²Chubu University Academy of Emerging Science, Kasugai, 487-0027, Japan

Rec: 2 June 2016, Rev: 7 December 2018, Accp: 18 December 2018,

Abstract

In this paper, we suggest Singular Monotone-singular monotone Newton Theorem-theorem and a relaxed Newton method for a nonlinear matrix equation $(F(X) = 0)$ where F is an order-convex matrix function. At first, we introduce Monotone-monotone Newton Theorem-theorem from Ortega and Rheinboldt (SIAM, 2000), and we show Singular Monotone-singular monotone Newton Theorem-theorem which shows that the Newton sequence converges to a solution of $(F(X) = 0)$ even if the Fréchet derivative at the solution is singular. At last, we provide a relaxed Newton method for $(F(X) = 0)$ which is better than the pure Newton method and we give some numerical experiments for the relaxed Newton method.

Keywords: Order-convex function, Nonlinear matrix equation, Newton method, Relaxed Newton method

Mathematics Subject Classification: 65H10

ArticleNote

Communicated by Carlos Conca.

Introduction

We introduce relevant definitions. A matrix A is called a *nonnegative (or positive) matrix* if all of entries of A are nonnegative (or positive) and we say that $(A \geq B)$ (or $(A > B)$) if $(A - B)$ is a nonnegative (or positive) matrix. We say that A and B are *comparable* if $(A \geq B)$ or $(A \leq B)$, and that A and B are *strictly comparable* if $(A > B)$ or $(A < B)$. A matrix C is called a *Z-matrix* if there exist a real number r and a nonnegative matrix N such that $(C = rI - N)$. C is called an *M-matrix* if $(r \geq \rho(N))$ where $(\rho(N))$ is the spectral radius of N . The linear operator $(\mathrm{vec}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{mn})$ is defined by

MediaObjects/40314_2019_1021_Equ1.gif

$$\begin{aligned} \mathrm{vec}(A) &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}^T \quad \text{for } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{aligned} \quad (1.1)$$

where (a_k) 's are column vectors of A .

In many applications, finding a solution of a nonlinear matrix equation is one of the important issues. Nonsymmetric algebraic Riccati equations occur in applied probability and transportation theory and have been extensively studied. See the details in Rogers (1994), Juang (1995), Benner and Byers (1998), Guo and Laub (2000), Guo et al. (2006), Guo and Higham (2007). Markovian binary trees are found in biology, epidemiology, and telecommunication systems (Haccou et al. 2005; Kimmel and Axelrod 2002; Hautphenne et al. 2006). Matrix polynomial equations occur in the theory of differential equations, system theory, network theory, stochastic theory, quasi-birth-and-death and other areas (Alfa 2003; Bean et al. 1997; Geoffrey et al. 1985; Gohberg et al. 1982; Lancaster 1966; Lancaster and Tismenetsky 1985; He and Neuts 2001; Dario 2005; Latouche and Ramaswami 1999).

These equations have the order-convex properties in common. Order-convex functions have been studied in Vandergraft (1967), Ortega and Rheinboldt (2000). Ortega and Rheinboldt (2000) showed that the convergence of the Newton method for a nonlinear equation $(F(X) = 0)$ if F is order \mathbb{R} -convex.

Davis (1981, 1983) and Higham and Kim (2000, 2001) considered the Newton method for a quadratic matrix equation. Guo (2001) considered a nonsymmetric algebraic Riccati equation, and they proposed iteration algorithms which converge to the minimal positive solution. In some different types of quadratic equations, (Kim 2008) showed that the minimal positive solutions also can be found by the Newton method with zero starting matrix. Hautphenne et al. (2008) studied the Newton method for the Markovian binary tree. Seo et al. (2013); Seo and Kim (2014) studied about the Newton iteration for a quadratic matrix equation and a matrix polynomial equation. In Seo and Kim (2014), they provided a relaxed Newton method whose convergence is faster than the pure one. The definition of the minimal nonnegative or positive solution is given as the following.

Definition 1.1

Let $(S_{\{1\}})$ and $(S_{\{2\}})$ be nonnegative solutions of a matrix equation $(F(X) = 0)$. Then, $(S_{\{1\}})$ is called the *minimal nonnegative solution* of $(F(X) = 0)$ and $(S_{\{2\}})$ is called the *maximal nonnegative solution* of $(F(X) = 0)$ if

$$S_{\{1\}} \leq S \leq S_{\{2\}}, \quad (1.2)$$

for any nonnegative solution S of $(F(X) = 0)$. Similarly, $(S_{\{1\}})$ and $(S_{\{2\}})$ are called the *minimal positive solution* and the *maximal positive solution*, respectively, if positive solutions $(S_{\{1\}})$ and $(S_{\{2\}})$ satisfy (1.2) for any positive solution S .

Some properties of nonnegative matrices and M -matrices are given in the following theorem.

Theorem 1.2

(Guo and Higham 2007, Theorem 2.1) (Poole and Boullion 1974, Theorem 2.1) For a Z -matrix A , the following are equivalent:

- (1) A is a nonsingular M -matrix.
- (2) (A^{-1}) is nonnegative.
- (3) $(Av > 0)$ for some vector $(v > 0)$.
- (4) All eigenvalues of A have positive real parts.

Theorem 1.3

(Roger 1985, Theorem 8.1.18) Let $(A, B \in \mathbb{R}^{n \times n})$. If $(|A| \leq B)$, then $(\rho(A) \leq \rho(|A|) \leq \rho(B))$ where $(|A|_{ij} = |A_{ij}|)$.

Theorem 1.4

(Guo and Higham 2007, Lemma 2.2) Let $(A \in \mathbb{R}^{m \times m})$ be a nonsingular M -matrix. If B is a Z -matrix and $(B \geq A)$, then B is also a nonsingular M -matrix.

The outline of the paper is as follows. In Section-Sect. 2, we introduce order-convex functions, and review its properties. In Section-Sect. 3, we introduce (Ortega and Rheinboldt 2000, Monotone monotone Newton Theoremtheorem) and provide a modified theorem which shows that the Newton sequence converges to a solution of $(F(X) = 0)$ where F is the order \mathbb{R} -convex even if the Fréchet derivative at the solution is singular. We provide a relaxed Newton method for an order-convex matrix equation in Section-Sect. 4. Numerical experiments are given in Section-Sect. 4.1.

Order-Convex Functionsconvex functions

For a matrix function, the Gâteaux and Fréchet derivative are defined as the followings:

Definition 2.1

(Higham [2008](#)) Let $(F : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{p \times q})$ be a matrix function. The *Gâteaux derivative* of F at X in the direction E is defined by

MediaObjects/40314_2019_1021_Equ3.gif

$$G_F(X, E) = \lim_{t \rightarrow 0} \frac{F(X + tE) - F(X)}{t}. \quad (2.1)$$

Assume that $(\|\cdot\|_m : \mathbb{C}^{m \times n} \rightarrow \mathbb{R})$ and $(\|\cdot\|_p : \mathbb{C}^{p \times q} \rightarrow \mathbb{R})$ are norms, and a linear transformation $(L_F(X) : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{p \times q})$ satisfies

MediaObjects/40314_2019_1021_Equ4.gif

$$F(X + H) - F(X) = L_F(X)(H) + o(H), \quad (2.2)$$

where $\frac{o(H)}{\|H\|_m} \rightarrow 0$ as $\|H\|_m \rightarrow 0$. Then, $(L_F(X)(H))$ is called the *Fréchet derivative* of F at X in the direction H .

If the Fréchet derivative of f exists at X , then it is equal to the Gâteaux derivative (Higham [2008](#)). In this paper, we deal with Fréchet differentiable functions. So, we denote the Gâteaux and Fréchet derivative of F at X in the direction H as $(F'[X](H))$. $(F'[X])$ is a linear operator from $(\mathbb{C}^{m \times n})$ to $(\mathbb{C}^{p \times q})$. Thus, $(\text{vec} \circ F'[X] \circ \text{vec}^{-1})$ is a $(pq \times mn)$ matrix. For convenience, we denote $(F'[X])$ this matrix.

The followings are the definition and an important theorem of order-convex functions.

Definition 2.2

(Vandergraft [1967](#); Ortega and Rheinboldt [2000](#)) Let $(F : \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q})$ be a function. Then, F is *order-convex* on a convex subset $(\mathcal{D}_0 \subset \mathcal{D})$ if

MediaObjects/40314_2019_1021_Equ5.gif

$$F(tX + (1-t)Y) \leq tF(X) + (1-t)F(Y), \quad (2.3)$$

where $(X, Y \in \mathcal{D}_0)$ are comparable and $(t \in (0, 1))$.

Theorem 2.3

(Ortega and Rheinboldt [2000](#), Theorem 13.3.2.) Let $(F : \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q})$ be Gâteaux differentiable on the convex set $(\mathcal{D}_0 \subset \mathcal{D})$. Then, the following statements are equivalent.

- (1) F is order-convex on (\mathcal{D}_0) ;
- (2) $(F(Y) - F(X)) \geq F'[X](Y - X)$ for comparable $(X, Y \in \mathcal{D}_0)$;
- (3) $(\left(F[Y] - F[X] \right) (Y - X) \geq 0)$ for comparable $(X, Y \in \mathcal{D}_0)$.

Let a matrix equation

MediaObjects/40314_2019_1021_Equ6.gif

$$F(X) = 0 \quad (2.4)$$

be given, where $(F : \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times m})$. If F is order-convex on $(\mathcal{D}_0 \subset \mathcal{D})$, then we say that [\(2.4\)](#) is an *order-convex matrix equation* on (\mathcal{D}_0) , for the convenience.

Consider a fixed point iteration

MediaObjects/40314_2019_1021_Equ7.gif

$$X_{i+1} = G(X_i), \quad (2.5)$$

where G is a Gâteaux differentiable and order-convex matrix function on (\mathcal{D}_0) .

Lemma 2.4

Let (X_i) be a sequence defined by [\(2.5\)](#) with a given $(X_0 \in \mathcal{D}_0)$. If (X_i)

and (X_l) are comparable and in $(\frac{D}{0})$ for all $(j, l \in \mathbb{N}_0)$ where $(\mathbb{N}_0 = \mathbb{N} \cup \{0\})$, then

MediaObjects/40314_2019_1021_Equ8.gif

$$\begin{aligned} G[X_{l-1}](X_{j-1} - X_{l-1}) \leq X_j - X_l \leq G[X_{j-1}](X_{j-1} - X_{l-1}). \end{aligned} \quad (2.6)$$

Furthermore, if (X_i) converges to $(\frac{D}{0})$, then

MediaObjects/40314_2019_1021_Equ9.gif

$$\begin{aligned} G[X_{j-1}](S - X_{j-1}) \leq S - X_j \leq G[S](S - X_{j-1}). \end{aligned} \quad (2.7)$$

Proof

Since (X_j) and (X_l) are comparable for all $(j, l \in \mathbb{N}_0)$, we obtain that

MediaObjects/40314_2019_1021_Equ10.gif

$$\begin{aligned} \begin{array}{rcl} X_j - X_l & \& = & G(X_{j-1}) - G(X_{l-1}) \\ G[X_{l-1}](X_{j-1} - X_{l-1}) & \& \leq & \end{array} \quad (2.8) \end{aligned}$$

by Theorem 2.3.

On the other hand, we obtain that

MediaObjects/40314_2019_1021_Equ11.gif

$$\begin{aligned} X_j - X_l = G(X_{j-1}) - G(X_{l-1}) \leq G[X_{j-1}](X_{j-1} - X_{l-1}). \end{aligned} \quad (2.9)$$

Therefore, (2.6) holds for all $(j, l \in \mathbb{N}_0)$.

Obviously, S is comparable with (X_j) for all $(j \in \mathbb{N}_0)$ and $(S = G(S))$. Thus,

MediaObjects/40314_2019_1021_Equ53.gif

$$\begin{aligned} G[X_{j-1}](S - X_{j-1}) \leq G(S) - G(X_{j-1}) \leq G[S](S - X_{j-1}). \end{aligned}$$

Since $(S - X_j) = G(S) - G(X_{j-1})$, (2.7) holds. \square

Theorem 2.5

Let (X_i) be a sequence defined in Lemma 2.4. If (X_i) is monotone nondecreasing, converges to $(\frac{D}{0})$, and $(G'[S] \geq 0)$, then

MediaObjects/40314_2019_1021_Equ12.gif

$$\begin{aligned} \limsup_i \rho \{X_i - S\} \leq \rho \{G'[S]\}. \end{aligned} \quad (2.10)$$

Proof

By Lemma 2.4, we get the inequality

MediaObjects/40314_2019_1021_Equ13.gif

$$\begin{aligned} \begin{array}{rcl} S - X_i & \& \leq & G[S](S - X_{i-1}) \\ G[S](S - X_{i-2}) & \& \leq & \dots \leq & G[S](S - X_0) \end{array} \quad (2.11) \end{aligned}$$

Applying 2-norm and Frobenius norm, we obtain that

MediaObjects/40314_2019_1021_Equ14.gif

$$\begin{aligned} \begin{array}{rcl} \|X_i - S\|_2 & \leq & \|G[S](S - X_{i-1})\|_2 \\ \|G[S](S - X_{i-2})\|_2 & \leq & \dots \leq & \|G[S](S - X_0)\|_2 \end{array} \quad (2.12) \end{aligned}$$

Therefore,

MediaObjects/40314_2019_1021_Equ54.gif

$$\begin{aligned} \begin{array}{rcl} \limsup_i \rho \{X_i - S\} & \leq & \limsup_i \rho \{G[S](S - X_{i-1})\} \\ \limsup_i \rho \{G[S](S - X_{i-2})\} & \leq & \dots \leq & \limsup_i \rho \{G[S](S - X_0)\} \end{array} \end{aligned}$$

by (Roger 1985, Corollary 5.6.14). \square

Corollary 2.6

Let (X_i) be a sequence in Lemma 2.4. Suppose that the following conditions hold.

- (1) $\{X_i\}$ is monotone nondecreasing and converges to S ,
- (2) $(S - X_0) > 0$,
- (3) $\{G[X_0]\}$ is a nonnegative matrix such that $\{G[X_0]\} \mathbf{1} > 0$,
- (4) $\{G[X_i]\}$ is monotone nondecreasing,
- (5) $\{G[\cdot]\}$ is continuous on $\{D\}$.

Then,

MediaObjects/40314_2019_1021_Equ15.gif

$$\limsup_{i \rightarrow \infty} \sqrt[i]{\|X_i - S\|} = \rho(G[S]). \quad (2.13)$$

Proof

Note that $\lim_{i \rightarrow \infty} G[X_i] = G[S]$. So, for any $\epsilon > 0$, there exists an integer N such that

MediaObjects/40314_2019_1021_Equ55.gif

$$\rho(G[X_N]) \leq \rho(G[S]) + \epsilon.$$

From the inequality

MediaObjects/40314_2019_1021_Equ16.gif

$$\begin{aligned} \|S - X_i\| &\leq \|S - X_{i-1}\| + \|G[X_{i-1}](S - X_{i-1})\| \\ &\leq \|S - X_{i-1}\| + \rho(G[X_{i-1}]) \|S - X_{i-1}\| \\ &\leq \|S - X_{i-1}\| + \rho(G[X_{i-1}]) \|S - X_{i-2}\| \\ &\leq \dots \\ &\leq \|S - X_0\|. \end{aligned} \quad (2.14)$$

we obtain that

MediaObjects/40314_2019_1021_Equ17.gif

$$\begin{aligned} \|S - X_i\| &\leq \|S - X_{i-1}\| + \rho(G[X_{i-1}]) \|S - X_{i-1}\| \\ &\leq \|S - X_{i-1}\| + \rho(G[X_{i-1}]) \|S - X_{i-2}\| \\ &\leq \dots \\ &\leq \|S - X_0\|. \end{aligned} \quad (2.15)$$

Since $\rho(G[X_0]) \leq \rho(G[S]) + \epsilon$, there exist $c > 0$ and a nonnegative vector v such that

MediaObjects/40314_2019_1021_Equ56.gif

$$\begin{aligned} \|S - X_i\| &\leq \|S - X_0\| + \rho(G[X_0]) \|S - X_0\| \\ &\leq \|S - X_0\| + \rho(G[X_0]) \|S - X_1\| \\ &\leq \dots \\ &\leq \|S - X_0\|. \end{aligned}$$

Therefore,

MediaObjects/40314_2019_1021_Equ57.gif

$$\begin{aligned} \limsup_{i \rightarrow \infty} \sqrt[i]{\|S - X_i\|} &\leq \limsup_{i \rightarrow \infty} \sqrt[i]{\|S - X_0\| + \rho(G[X_0]) \|S - X_0\|} \\ &\leq \limsup_{i \rightarrow \infty} \sqrt[i]{\|S - X_0\|} \\ &= \rho(G[S]) + \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have

MediaObjects/40314_2019_1021_Equ58.gif

$$\limsup_{i \rightarrow \infty} \sqrt[i]{\|S - X_i\|} = \rho(G[S]).$$

by Theorem 2.5. \square

For example, let the fixed point iteration

MediaObjects/40314_2019_1021_Equ18.gif

$$X_{i+1} = L^{-1}(X_i C X_i + X_i D + A X_i + B) \quad (2.16)$$

be given, where $A, D \geq 0$, $(B, C > 0)$ and $L(X) = A_1 X + X D_1$ such that

(A_{-1}) and (D_{-1}) are M -matrices. If (X_{-0}) is the solution of the equation $(A_{-1}X_{-0} + X_{-0}D_{-1} = B)$, and we define a matrix function G by

MediaObjects/40314_2019_1021_Equ19.gif

$$G(X) = \mathcal{L}^{-1}(XCX + XD_{-2} + A_{-2}X + B), \quad (2.17)$$

then we can easily verify that G is order \square convex on $(\mathbb{R}^{+})^m \times \mathbb{R}^n := \{X \in \mathbb{R}^{m \times n} \mid X \geq 0\}$, and (X_{-i}) satisfies the conditions of Corollary 2.6. Therefore,

MediaObjects/40314_2019_1021_Equ20.gif

$$\limsup_i \sqrt[i]{\|G'(S)\|} = \rho(G'(S)), \quad (2.18)$$

where $(G'(S))$ is the derivative of G at S and $\mathcal{G}(S) = \mathrm{vec}(\mathrm{vec}^{\circ} G'(S) \mathrm{vec}^{\circ -1})$. In (Guo and Laub 2000, Theorem 3.2), Guo and Laub also proved (2.18) with direct calculations.

The Newton Method method for an Orderorder-Convex Functionconvex function

The Newton method for solving a matrix equation $(F(X) = 0)$ is defined by a given initial matrix (X_{-0}) and the iteration

MediaObjects/40314_2019_1021_Equ21.gif

$$X_{i+1} = X_i - F'(X_i)^{-1}(F(X_i)), \quad i = 0, 1, 2, \dots, \quad (3.1)$$

where $(F: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n})$ is a matrix function and $(F'(X))$ is the Fréchet derivative of F at X . If $(F'(X_i))$ is nonsingular, then (3.1) is equivalent to

MediaObjects/40314_2019_1021_Equ22.gif

$$\{ \left\{ \begin{array}{l} -F'(X_i)(H_i) = F(X_i), \\ X_{i+1} = X_i + H_i, \end{array} \right. \}_{i=0,1,2,\dots}. \quad (3.2)$$

To get (H_i) in the equation $(-F'(X_i)(H_i) = F(X_i))$, we apply the operator (vec) in both sides of the equation. Then, we obtain the following equation:

MediaObjects/40314_2019_1021_Equ23.gif

$$-\mathcal{F}'[X_i] \mathrm{vec}(H_i) = \mathrm{vec}(\mathrm{vec}^{\circ} F(X_i) \mathrm{vec}^{\circ}). \quad (3.3)$$

Since $(-\mathcal{F}'[X_i])$ is a nonsingular $(mn \times mn)$ matrix, we can obtain the unique (H_i) for each $(i = 0, 1, 2, \dots)$.

Theorem 3.1

(Ortega and Rheinboldt 2000, Monotone monotone Newton Theoremtheorem) Let $(F: \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n})$, and assume that there exist $(X_{-0}, Y_{-0} \in \mathcal{D})$ such that

MediaObjects/40314_2019_1021_Equ24.gif

$$X_{-0} \leq Y_{-0}, \quad [X_{-0}, Y_{-0}] \subset \mathcal{D}, \quad F(X_{-0}) \leq 0 \leq F(Y_{-0}), \quad (3.4)$$

where $([X_{-0}, Y_{-0}] = \{A \mid X_{-0} \leq A \leq Y_{-0}\})$.

Suppose that F is continuous, Gâteaux differentiable, order \square convex on $([X_{-0}, Y_{-0}])$, and that, for each $(X \in [X_{-0}, Y_{-0}])$, $(\mathcal{F}'[X]^{-1})$ exists and is nonnegative. Then, the Newton iteration defined by

MediaObjects/40314_2019_1021_Equ25.gif

$$Y_{k+1} = Y_k - F'(Y_k)^{-1}(F(Y_k)), \quad k = 0, 1, 2, \dots, \quad (3.5)$$

converges to $(S \in [X_{-0}, Y_{-0}])$ and is monotone nonincreasing. Furthermore, if $(\mathcal{F}'[\cdot] \cdot \dot{})$ is either continuous at S or monotone nondecreasing on $([X_{-0}, Y_{-0}])$, then S is the unique solution of $(F(X) = 0)$ in $([X_{-0}, Y_{-0}])$.

Corollary 3.2

Let $(F: \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n})$, and suppose that there exist $(X_{-0}, Y_{-0} \in \mathcal{D})$ such that

MediaObjects/40314_2019_1021_Equ26.gif

$$\begin{aligned} X_0 \leq Y_0, \quad [X_0, Y_0] \subset \mathfrak{D}, \quad \{ F(X_0) \\ \geq 0 \geq F(Y_0) \}. \end{aligned} \quad (3.6)$$

Suppose that F is continuous, Gâteaux differentiable, order α convex on $([X_0, Y_0])$, and that, for each $(X \in [X_0, Y_0])$, $(-F'[X]^{-1})$ exists and is nonnegative. Then, the Newton iteration

MediaObjects/40314_2019_1021_Equ27.gif

$$X_{i+1} = X_i - F'[X_i]^{-1}(F(X_i)), \quad \text{quad } i = 0, 1, 2, \dots, \quad (3.7)$$

converges to $(S \in [X_0, Y_0])$ and is monotone nondecreasing. Furthermore, if $(F'[\cdot] \cdot \cdot \cdot)$ is either continuous at S or monotone nonincreasing on $([X_0, Y_0])$, then S is the unique solution of $(F(X) = 0)$ in $([X_0, Y_0])$.

Proof

Define $(F_1 : [X_0, Y_0] \rightarrow [X_0, Y_0])$ by $(F_1(X) = X_0 + Y_0 - X)$ and $(F_2(X) = (F \circ F_1)(X))$. Since (F_1) is Fréchet differentiable in $([X_0, Y_0])$, (F_2) is Gâteaux differentiable $([X_0, Y_0])$ in and

MediaObjects/40314_2019_1021_Equ59.gif

$$F'_2[X](H) = \left(F'[F_1(X)] \circ F'_1[X] \right) (H) = -F'[X_0 + Y_0 - X](H),$$

by (Ortega and Rheinboldt 2000, 3.1.7. Chain Rule.). Furthermore, for comparable $(X, Y \in [X_0, Y_0])$,

MediaObjects/40314_2019_1021_Equ60.gif

$$F_2(Y) - F_2(X) = F(X_0 + Y_0 - Y) - F(X_0 + Y_0 - X) \geq F[X_0 + Y_0 - X](X_0 + Y_0 - Y - X_0 - Y_0 + X) = -F[X_0 + Y_0 - X](Y - X) = F'_2[X](Y - X).$$

Since $(-F'[X]^{-1})$ exists and is nonnegative for all $(X \in [X_0, Y_0])$, $(F'_2[X]^{-1} = -F'[X_0 + Y_0 - X]^{-1})$ exists and is nonnegative for all $(X \in [X_0, Y_0])$. At last,

MediaObjects/40314_2019_1021_Equ61.gif

$$F_2(X_0) = F(Y_0) \leq 0 \leq F(X_0) = F_2(Y_0).$$

Therefore, by Theorem 3.1, the Newton iteration

MediaObjects/40314_2019_1021_Equ62.gif

$$Y_{i+1} = Y_i - F'_2[Y_i]^{-1}(F_2(Y_i)), \quad \text{quad } i = 0, 1, 2, \dots,$$

converges to $(S_2 \in [X_0, Y_0])$ and is monotone nonincreasing.

Put $(X_k = X_0 + Y_0 - Y_k)$. Then,

MediaObjects/40314_2019_1021_Equ63.gif

$$X_{i+1} = X_0 + Y_0 - Y_{i+1} = X_0 + Y_0 - Y_i + F'_2[Y_i]^{-1}(F_2(Y_i)) = X_i - F'[X_0 + Y_0 - Y_i]^{-1}(F(X_0 + Y_0 - Y_i)) = X_i - F'[X_i]^{-1}(F(X_i)).$$

So, (X_i) is also the Newton sequence, and monotone nondecreasing, and converges to $(S = X_0 + Y_0 - S_2)$.

Finally, since $(F'[X] = -F'_2[X_0 + Y_0 - X])$, $(F'[\cdot] \cdot \cdot \cdot)$ is monotone nonincreasing on $([X_0, Y_0])$ if and only if $(F'_2[\cdot] \cdot \cdot \cdot)$ is monotone nondecreasing on $([X_0, Y_0])$. Moreover, $(F'[\cdot] \cdot \cdot \cdot)$ is continuous at S if and only if $(F'_2[\cdot] \cdot \cdot \cdot)$ is continuous at (S_2) since $(\|X - S\| = \|X - X_0 - Y_0 + S_2\| = \|X_0 + Y_0 - X - S_2\|)$ and

MediaObjects/40314_2019_1021_Equ64.gif

$$\|F'[X] - F'[S]\| = \|F'_2[X_0 + Y_0 - X] - F'_2[X_0 + Y_0 - S]\| = \|F'_2[X_0 + Y_0 - X] - F'_2[S_2]\|$$

for any norm $(\|\cdot\|)$.

Suppose that a matrix function F and a matrix $(X_0 = 0)$ satisfies satisfy the hypothesis of Corollary 3.2. Then, the corollary shows that the Newton sequence converges to a nonnegative solution S . Moreover, if S is the unique solution in $([X_0, Y_0])$, then S is the minimal nonnegative solution. All of these are based on the hypothesis that $(F'[S])$ is nonsingular. But, sometimes, the derivative is singular.

For example, let $(f : \mathbb{R}^2 \rightarrow \mathbb{R}^2)$ be defined by

MediaObjects/40314_2019_1021_Equ28.gif

$$\left(\begin{aligned} f(\mathbf{x}) &= \begin{bmatrix} (x_1-1)^2 + (x_2-1)^2 \\ \frac{1}{2}(x_1-1)^2 + (x_2-1)^2 \end{bmatrix}, \quad \text{where } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned} \right) \quad (3.8)$$

Then, f is differentiable and order $\frac{1}{2}$ convex on (\mathbb{R}^2) , $(f(\mathbf{x}) \geq 0)$ for all $(\mathbf{x} \in \mathbb{R}^2)$, and

MediaObjects/40314_2019_1021_Equ65.gif

$$\left(\begin{aligned} f'(\mathbf{x}) &= \begin{bmatrix} 2x_1 - 2 & 2x_2 - 2 \end{bmatrix} \quad \text{if } x_1 \neq 1 \text{ and } x_2 \neq 1 \\ &= \begin{bmatrix} 0 & 0 \end{bmatrix} \quad \text{if } x_1 = 1 \text{ and } x_2 = 1 \end{aligned} \right)$$

If $(f(\mathbf{x}_0) = 0)$ and $(\mathbf{y}_0 = [1 \ 1]^T)$, then $(f(\mathbf{y}_0) = 0)$ and $(f'(\mathbf{y}_0) = 0)$ is singular. So, f satisfies the conditions of Corollary 3.2 except that $(-f'(\mathbf{y}_0)^{-1})$ exists. But, the Newton sequence (\mathbf{x}_i) converges to $([1 \ 1]^T)$.

So, we provide a modified Monotone monotone Newton Theorem theorem for the case that the derivative at S is singular. Before we do that, we introduce the definition of strictly order $\frac{1}{2}$ convex and relevant properties.

Definition 3.3

Let $(F : \mathcal{D} \subset \mathbb{R}^m \rightarrow \mathbb{R}^p)$ be a function. Then, F is *strictly order $\frac{1}{2}$ convex* on a convex subset $(\mathcal{D}_0 \subset \mathcal{D})$ if

MediaObjects/40314_2019_1021_Equ29.gif

$$\left(F(tX + (1-t)Y) < tF(X) + (1-t)F(Y), \quad \text{for } t \in (0,1) \right) \quad (3.9)$$

where $(X, Y \in \mathcal{D}_0)$ are strictly comparable and $(t \in (0,1))$.

Theorem 3.4

(cf. (Ortega and Rheinboldt 2000, Theorem 3.2.2)) Suppose that $(f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^1)$ is Gâteaux differentiable at each point of a convex set $(D_0 \subset D)$. Then, for any two points $(x, y \in D_0)$, there is a $(t \in (0,1))$ such that

MediaObjects/40314_2019_1021_Equ66.gif

$$\left(f(y) - f(x) = f'[x + t(y-x)](y-x), \quad \text{for } t \in (0,1) \right)$$

Theorem 3.5

Let (\mathcal{D}) be open, $(\mathcal{D}_0 \subset \mathcal{D})$ be a convex set, and $(F : \mathcal{D} \subset \mathbb{R}^m \rightarrow \mathbb{R}^p)$ be Gâteaux differentiable on (\mathcal{D}) . Then, the following statements are equivalent.

- (1) F is strictly order $\frac{1}{2}$ convex on (\mathcal{D}_0) ;
- (2) $(F(Y) - F(X) > F'[X](Y - X))$ for strictly comparable $(X, Y \in \mathcal{D}_0)$;
- (3) $([F'(Y) - F'(X)](Y - X) > 0)$ for strictly comparable $(X, Y \in \mathcal{D}_0)$.

Proof

Let $(X, Y \in \mathcal{D}_0)$ be strictly comparable.

Suppose that (2) holds. Let $(Z = tX + (1-t)Y)$ for some $(t \in (0,1))$. Then, Z is strictly comparable with X and Y . So, we obtain that

MediaObjects/40314_2019_1021_Equ30.gif

$$\left(F(X) - F(Z) > F'[Z](X - Z), \quad \text{for } t \in (0,1) \right) \quad (3.10)$$

MediaObjects/40314_2019_1021_Equ31.gif

$$\left(F(Y) - F(Z) > F'[Z](Y - Z), \quad \text{for } t \in (0,1) \right) \quad (3.11)$$

Multiplying t to (3.10) and $(1-t)$ to (3.11), and adding these, we obtain that

MediaObjects/40314_2019_1021_Equ32.gif

$$[\begin{aligned} tF(X) + (1-t)F(Y) - F(Z) > F[Z](tX + (1-t)Y - Z) = 0, \end{aligned}] \quad (3.12)$$

i.e., (1) holds.

Conversely, suppose that (1) holds. Let $Z_{\{1\}} = (1 - t_{\{1\}})X + t_{\{1\}}Y$ for some $t_{\{1\}} \in (0,1)$. Then,

MediaObjects/40314_2019_1021_Equ67.gif

$$[\begin{aligned} \begin{array}{l} (1 - t_{\{1\}})F(X) + t_{\{1\}}F(Y) > F(Z_{\{1\}}) \\ \left[F(Y) - F(X) \right] > F(Z_{\{1\}}) - F(X) \end{array} \\ \Leftrightarrow F(Y) - F(X) > (1/t_{\{1\}}) \left[F(Z_{\{1\}}) - F(X) \right]. \end{aligned}]$$

Since X and $(X + t_{\{2\}}(Z_{\{1\}} - X))$ are strictly comparable for any $t_{\{2\}} \in (0,1)$, we obtain that

MediaObjects/40314_2019_1021_Equ68.gif

$$[\begin{aligned} F(Z_{\{1\}}) - F(X) > \frac{1}{t_{\{2\}}} \left[F(X + t_{\{2\}}(Z_{\{1\}} - X)) - F(X) \right]. \end{aligned}]$$

In the view of Gâteaux differentiable of F ,

MediaObjects/40314_2019_1021_Equ69.gif

$$[\begin{aligned} F(Z_{\{1\}}) - F(X) \geq F'[X](Z_{\{1\}} - X) \end{aligned}]$$

follows as $t_{\{2\}} \rightarrow 0$.

By the way, $Z_{\{1\}} - X = t_{\{1\}}(Y - X)$. Therefore,

MediaObjects/40314_2019_1021_Equ70.gif

$$[\begin{aligned} F(Y) - F(X) > \frac{1}{t_{\{1\}}} \left[F(Z_{\{1\}}) - F(X) \right] \geq \frac{1}{t_{\{1\}}} F'[X](Z_{\{1\}} - X) = F'[X](Y - X). \end{aligned}]$$

Suppose that (2) holds. Then, adding the following two inequalities:

MediaObjects/40314_2019_1021_Equ71.gif

$$[\begin{aligned} F(Y) - F(X) > F'[X](Y - X) \quad \text{and} \quad F(X) - F(Y) > F'[Y](X - Y), \end{aligned}]$$

we easily obtain (3).

Conversely, if (3) holds, then

MediaObjects/40314_2019_1021_Equ72.gif

$$[\begin{aligned} (F'_{\{ij\}}[Y] - F'_{\{ij\}}[X])(Y - X) > 0, \quad i = 1, \dots, p, \quad j = 1, \dots, q, \end{aligned}]$$

where $(F'_{\{ij\}})$ is the (i, j) -entry function of F . By Theorem 3.4, there is $t_{\{ij\}} \in (0,1)$ such that

MediaObjects/40314_2019_1021_Equ73.gif

$$[\begin{aligned} F_{\{ij\}}(Y) - F_{\{ij\}}(X) = F'_{\{ij\}}[Z^{\{ij\}}](Y - X), \end{aligned}]$$

where $Z^{\{ij\}} = X + t_{\{ij\}}(Y - X)$. Since $Z^{\{ij\}}$ is strictly comparable with X and Y ,

MediaObjects/40314_2019_1021_Equ74.gif

$$[\begin{aligned} (F'_{\{ij\}}[Z^{\{ij\}}] - F'_{\{ij\}}[X])(Y - X) = (1/t_{\{ij\}})(F'_{\{ij\}}[Z^{\{ij\}}] - F'_{\{ij\}}[X])(Z^{\{ij\}} - X) > 0, \end{aligned}]$$

and, so,

MediaObjects/40314_2019_1021_Equ75.gif

$$[\begin{aligned} F_{\{ij\}}(Y) - F_{\{ij\}}(X) = F'_{\{ij\}}[Z^{\{ij\}}](Y - X) > F'_{\{ij\}}[X](Y - X), \end{aligned}]$$

for all $(i = 1, \dots, p)$ and $(j = 1, \dots, q)$. \square

Lemma 3.6

The following statements hold:

- (1) Let A and B be nonnegative matrices. If $(A \leq B)$ and $(B - A)$ is irreducible, then $(\rho(A) < \rho(B))$.
- (2) Let $(C \in \mathbb{R}^{n \times n})$ be a singular M -matrix. If D is a Z -matrix and $(D - C)$ is nonnegative and irreducible, then D is a nonsingular M -matrix.

Proof

To prove (1), suppose that $(0 \leq A \leq B)$, $(B - A)$ is irreducible, and $(\rho(A) = \rho(B))$. By Perron-Frobenius Theorem, there exists a vector $(x > 0)$ such that $((B - A)x = \rho(B - A)x)$ and $(\rho(B - A) > 0)$. Clearly, $(\rho(B) < \rho(B) + \rho(B - A))$. By (Roger 1985, Theorem 8.3.2), we obtain that $(By < \left[\rho(B) + \rho(B - A) \right] y)$ for any nonzero $(y \geq 0)$. Therefore,

MediaObjects/40314_2019_1021_Equ76.gif

(Singular **Monotone** ~~monotone~~ Newton Theorem) Let $(\mathcal{D} \subset \mathbb{R}^{m \times n})$ be open and $(F : \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n})$, and suppose that there exist $(X_0, Y_0) \in \mathcal{D}$ such that

MediaObjects/40314_2019_1021_Equ37.gif

$$[\begin{aligned} X_0 < Y_0, \quad [X_0, Y_0] \subset \mathcal{D}, \quad \{ F(X_0) > 0 \geq F(Y_0) \}. \end{aligned}] \quad (3.17)$$

Suppose that F is Fréchet differentiable on $([X_0, Y_0])$, $(A < B)$ implies that $(\mathcal{F}[B] - \mathcal{F}[A])$ is nonnegative and irreducible for all $(A, B \in [X_0, Y_0])$, $(-\mathcal{F}'[X_0])$ is a Z -matrix, and $(-\mathcal{F}'[Y_0])$ is a singular or nonsingular M -matrix. Then, the Newton iteration

MediaObjects/40314_2019_1021_Equ38.gif

$$[\begin{aligned} X_{i+1} = X_i - F'[X_i]^{-1}(F(X_i)), \quad i = 0, 1, 2, \dots, \\ \end{aligned}] \quad (3.18)$$

converges to $(S \in [X_0, Y_0])$ and is monotone increasing.

Relaxed Newton ~~Method~~ **method for Order-order-Convex Equations** **convex equations**

Benner and Byers (1998) provided an exact line search method along the Newton direction for solving generalized continuous-time algebraic Riccati equations. Higham and Kim (2001) improved Newton's method for solving quadratic matrix equations with line searches. Seo and Ki (2008) also provided Newton's method with exact line searches for the matrix polynomial equations. The idea of exact line search methods for matrix equations is to find (λ) which minimizes the merit function

MediaObjects/40314_2019_1021_Equ39.gif

$$[\begin{aligned} p(\lambda) = \| F(X_i) + \lambda H_i \|^2. \end{aligned}] \quad (4.1)$$

But, we may not apply the method to order-convex equations since we can not guarantee that $(\mathcal{F}'[X_i] + \lambda H_i)^{-1}$ exists and is nonnegative. By the way, (Seo and Kim 2014) provided the relaxed Newton method for a matrix polynomial equation(MPE) arising in stochastic models. The MPE under the condition in Seo and Kim (2014) is also an order-convex matrix equation on $(\{X \in \mathbb{R}^{m \times m} : X \geq 0\})$. So, in this section, we extend the proof of the relaxed Newton method for the order-convex matrix equations.

Theorem 4.1

Let $(F : \mathcal{D} \subset \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n})$, and suppose that there exist $(Y_0) \in \mathcal{D}$ such that

MediaObjects/40314_2019_1021_Equ40.gif

$$[\begin{aligned} X_0 \leq Y_0, \quad [X_0, Y_0] \subset \mathcal{D}, \quad F(X_0) \geq 0 \geq F(Y_0). \end{aligned}] \quad (4.2)$$

Suppose that F is continuous, Gâteaux differentiable, order α convex on $([X_0, Y_0])$, $(\mathcal{F}'[\cdot])$ is monotone nondecreasing on $([X_0, Y_0])$, and that, for each $(X \in [X_0, Y_0])$, $(-\mathcal{F}'[X]^{-1})$ exists and is nonnegative. Put (X_1) be the next Newton step for F of (X_0) . If we define that

MediaObjects/40314_2019_1021_Equ41.gif

$$[\begin{aligned} X_\lambda = X_0 - \lambda F'[X_0]^{-1}(F(X_0)), \end{aligned}] \quad (4.3)$$

where

MediaObjects/40314_2019_1021_Equ42.gif

$$[\begin{aligned} \lambda = \min \left\{ \left| \frac{[F(X_0)]_{jk} + [F(X_1)]_{jk}}{[F(X_0)]_{jk}} \right| : [F(X_0)]_{jk} > 0, \sim 1 \leq j \leq m, \sim 1 \leq k \leq n \right\} \\ , \quad \text{if } [F(X_0)]_{jk} = 0 \end{aligned}] \quad (4.4)$$

then $(X_1 \leq X_\lambda \leq Y_0)$ and $(F(X_\lambda) \geq 0)$.

Proof

From Corollary 3.2, $(\{X_i\})$ is well defined and monotone nondecreasing and we can easily check that $(F(X_i) \geq 0)$ for $(i = 1, 2, \dots)$. Thus, it is clear that $(\lambda \geq 1)$. From $(\lambda \geq 1)$,

MediaObjects/40314_2019_1021_Equ43.gif

$$[\begin{aligned} X_1 = X_0 - F'[X_0]^{-1}(F(X_0)) \leq X_0 - \lambda F'[X_0]^{-1}(F(X_0)) \end{aligned}]$$

$$F(X_{\lambda}) = X_{\lambda}. \quad (4.5)$$

On the other hand, since $F'(\cdot)$ is monotone nondecreasing on $[X_0, Y_0]$,

MediaObjects/40314_2019_1021_Equ79.gif

$$\begin{aligned} -F'(X_0)(X_2 - X_{\lambda}) &= -F'(X_0)\left(X_1 - F'(X_1)^{-1}(F(X_1)) - X_{\lambda}\right) \\ &= -F'(X_0)\left((\lambda - 1)F'(X_0)^{-1}(F(X_0)) - F'(X_1)^{-1}(F(X_1))\right) \\ &= (1 - \lambda)F(X_0) + F'(X_0)\left(F'(X_1)^{-1}(F(X_1))\right) \\ &= (1 - \lambda)F(X_0) + F'(X_1)\left(F'(X_1)^{-1}(F(X_1))\right) \\ &= F(X_0) + F(X_1) - \lambda F(X_0) \geq 0, \end{aligned}$$

where X_2 is the next Newton step of X_1 . Since $F'(X_0)^{-1}$ is nonnegative, $X_{\lambda} \leq X_2$. So, $X_{\lambda} \leq Y_0$. Since $X_{\lambda} \in [X_0, Y_0]$ and $F'(\cdot)$ is monotone nondecreasing on $[X_0, Y_0]$,

MediaObjects/40314_2019_1021_Equ80.gif

$$\begin{aligned} F(X_{\lambda}) &\geq F(X_1) + F'(X_1)(X_{\lambda} - X_1) \\ &= F(X_1) + (\lambda - 1)F'(X_1)\left(-F'(X_0)^{-1}(F(X_0))\right) \\ &\geq F(X_1) + (\lambda - 1)F'(X_0)\left(-F'(X_0)^{-1}(F(X_0))\right) \\ &= F(X_1) + F(X_0) - \lambda F(X_0) \geq 0. \end{aligned}$$

\square

Theorem 4.2

Let $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$, and suppose that there exist $Y_0 \in \mathbb{R}^m$ such that

MediaObjects/40314_2019_1021_Equ44.gif

$$X_0 \leq Y_0, \quad [X_0, Y_0] \subset \mathbb{R}^m, \quad F(X_0) \geq 0 \geq F(Y_0). \quad (4.6)$$

Suppose that F is continuous, Gateaux differentiable, order α convex on $[X_0, Y_0]$, $F'(\cdot)$ is monotone nondecreasing on $[X_0, Y_0]$, and that, for each $X \in [X_0, Y_0]$, $F'(X)^{-1}$ exists and is nonnegative. Then, for the relaxed Newton iteration

MediaObjects/40314_2019_1021_Equ45.gif

$$X_{i+1} = X_i - \lambda_i F'(X_i)^{-1}(F(X_i)), \quad (4.7)$$

the sequence $\{X_i\}$ is well defined, and monotone nondecreasing, and converges to $S \in [X_0, Y_0]$ where

MediaObjects/40314_2019_1021_Equ81.gif

$$\begin{aligned} \lambda_i &= \min \left\{ \left| \frac{F(X_{i+1}) - F(X_i)}{F(X_i)} \right| : \left| \frac{F(X_{i+1}) - F(X_i)}{F(X_i)} \right| > 0, \sim 1 \leq i \leq m, \sim 1 \leq k \leq n \right\}. \end{aligned}$$

Numerical Experiments

In this subsection, we give numerical experiments to compare the relaxed Newton method with the pure one. The tolerance of the algorithms is 10^{-16} and we stop the algorithms if the relative residuals are less than the tolerance. The relative residuals are given by

MediaObjects/40314_2019_1021_Equ46.gif

$$\begin{aligned} \frac{\|R(X_i)\|}{\|F\|} &= \frac{\|C\| \|X_i\| \|F\|^2 + \|D\| \|X_i\| \|F\| + \|A\| \|X_i\| \|F\| + \|B\| \|F\|}{\|F\|^2} \\ &= \frac{\|C\| \|X_i\| + \|D\| \|X_i\| + \|A\| \|X_i\| + \|B\|}{\|F\|}, \end{aligned} \quad (4.8)$$

in Example 4.3, and

MediaObjects/40314_2019_1021_Equ47.gif

$$\frac{\|M(x_i)\|}{\|F\|} = \frac{\|x_i\| \|F\| + \|a\| \|F\| + \|B\|}{\|F\|^2}, \quad (4.9)$$

in Example 4.4. The following experiments were executed by Matlab R2015a (8.5.0.197613) 64-bit with Celeron dual processor at 1.80GHz with 4GB of RAM.

Example 4.3

We consider the nonsymmetric algebraic Riccati equation

MediaObjects/40314_2019_1021_Equ48.gif

$$R(X) = CX - XD - AX + B = 0, \quad (4.10)$$

Table 1 The numbers of iterations for Example 4.3

(δ)	Pure	Relaxed	(δ)	Pure	Relaxed
(10^{-1})	6	5	(10^{-9})	19	13
(10^{-2})	8	6	(10^{-10})	20	14
(10^{-3})	10	7	(10^{-11})	21	15
(10^{-4})	11	8	(10^{-12})	23	16
(10^{-5})	13	9	(10^{-13})	24	17
(10^{-6})	14	10	(10^{-14})	24	18
(10^{-7})	16	11	(10^{-15})	25	18
(10^{-8})	17	12	(10^{-16})	25	18

where the matrix

MediaObjects/40314_2019_1021_Equ49.gif

$$\begin{aligned} K = \left[\begin{matrix} D & \\ & -C \end{matrix} \right] - B \quad A \end{aligned} \quad (4.11)$$

is a nonsingular or singular irreducible M -matrix. Guo (2001) proved the existence of the minimal nonnegative solution of (4.10) with (4.11). Xue et al. (2012) called this equation as an M -matrix algebraic Riccati equation(MARE). Let A , B , C , and D be given by

MediaObjects/40314_2019_1021_Equ50.gif

$$\begin{aligned} A = I_3 - (1-\delta) \begin{bmatrix} \frac{3}{8} \\ \frac{1}{20} \\ \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \quad B = \frac{3 + \delta}{3} \begin{bmatrix} \frac{1}{20} \\ \frac{1}{5} \\ \frac{3}{16} \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \quad C = (1-\delta) \begin{bmatrix} \frac{1}{12} \\ \frac{1}{16} \\ \frac{3}{16} \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \quad D = I_5 - \frac{3 + \delta}{3} \begin{bmatrix} \frac{1}{16} \\ \frac{1}{16} \\ \frac{1}{16} \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \end{bmatrix} \end{aligned} \quad (4.12)$$

for $(0 \leq \delta < 1)$, where $(\mathbf{1}_m \times n)$ is an $(m \times n)$ matrix whose all of elements equal to 1. For this case, K is a singular irreducible M -matrix and (4.10) is also an order-convex matrix equation. If $(\delta = 0)$, then the minimal nonnegative solution S of (4.10) equals to $(\frac{1}{5} \mathbf{1}_3 \times 5)$ and the Fréchet derivative at S is singular.

Table 1 reports the numbers of the pure and relaxed Newton iterations for given (δ) . Figure 1 shows the relative residuals of the iterations when $(\delta = 10^{-4})$ and $(\delta = 10^{-16})$, respectively. From them, we see that the relaxed Newton iteration is less than the pure one.

40314_2019_1021_Fig1_print.png

Fig. 1 The convergence of the relaxed and pure Newton methods in Example 4.3

Table 2 and Figure 2 show the average of total computation time and its efficiency of the pure and relaxed Newton methods for 300 trials with given (δ) . The results show that the relaxed Newton method reduces computation time about 15–20% than the pure one.

Table 1 The numbers of iterations for Example 4.3

(δ)	Pure	Relaxed	(δ)	Pure	Relaxed
(10^{-1})	65	19	(10^{-2})	86	20
(10^{-3})	107	21	(10^{-4})	118	23
(10^{-5})	139	24	(10^{-6})	141	24
(10^{-7})	161	25	(10^{-8})	171	25
(10^{-9})	191	13	(10^{-10})	201	14
(10^{-11})	211	15	(10^{-12})	231	16
(10^{-13})	241	17	(10^{-14})	241	18
(10^{-15})	251	18	(10^{-16})	251	18

40314_2019_1021_Fig1_print.png

Fig. 1

The

convergences of the relaxed and pure Newton methods in Example 4.3

Table 2 The averages of total computation time for 300 trials in Example 4.3

(δ)	Pure	Relaxed	Efficiency (%)	(δ)	Pure	Relaxed	Efficiency (%)
------------	------	---------	----------------	------------	------	---------	----------------

$\backslash(10^{\{-1\}}\backslash)$	0.00205	0.00197	3.91%91	$\backslash(10^{\{-9\}}\backslash)$	0.00706	0.00571	23.64%64
$\backslash(10^{\{-2\}}\backslash)$	0.00291	0.00252	15.37%37	$\backslash(10^{\{-10\}}\backslash)$	0.00744	0.00618	20.44%44
$\backslash(10^{\{-3\}}\backslash)$	0.00363	0.00295	23.45%15	$\backslash(10^{\{-11\}}\backslash)$	0.00784	0.00670	16.94%94
$\backslash(10^{\{-4\}}\backslash)$	0.00391	0.00337	16.54%54	$\backslash(10^{\{-12\}}\backslash)$	0.00863	0.00716	20.59%59
$\backslash(10^{\{-5\}}\backslash)$	0.00467	0.00382	22.26%26	$\backslash(10^{\{-13\}}\backslash)$	0.00898	0.00765	17.39%39
$\backslash(10^{\{-6\}}\backslash)$	0.00510	0.00431	18.20%20	$\backslash(10^{\{-14\}}\backslash)$	0.00902	0.00810	11.33%33
$\backslash(10^{\{-7\}}\backslash)$	0.00587	0.00477	23.21%21	$\backslash(10^{\{-15\}}\backslash)$	0.00941	0.00815	15.58%58
$\backslash(10^{\{-8\}}\backslash)$	0.00627	0.00525	19.37%37	$\backslash(10^{\{-16\}}\backslash)$	0.00940	0.00810	16.42%12

40314_2019_1021_Fig2_print.png

Fig. 2 The time efficiency of the relaxed Newton method in Example 4.3

Example 4.4

Let a Markovian binary tree equation(MBTE)

MediaObjects/40314_2019_1021_Equ51.gif

$$\backslash[\backslash\begin{aligned} M(x) = B(x \otimes x) - x + a = 0 \end{aligned}\backslash] \quad (4.13)$$

be given, where $\backslash(a \in \mathbb{R}^n\backslash)$ and $\backslash(B \in \mathbb{R}^{n \times n^2}\backslash)$ such that $\backslash(0 \leq a \leq 1\backslash)$, $\backslash(a \neq 0\backslash)$, $\backslash(B \geq 0\backslash)$, and $\backslash(a + B \mathbf{1}_- = \mathbf{1}_-\backslash)$. The MBTE is called subcritical, supercritical, or critical if $\backslash(\rho(M)\backslash)$ of the nonnegative matrix

MediaObjects/40314_2019_1021_Equ52.gif

$$\backslash[\backslash\begin{aligned} M = B(\mathbf{1}_- \otimes I + I \otimes \mathbf{1}_-) = B(\mathbf{1}_- \oplus \mathbf{1}_-) \end{aligned}\backslash] \quad (4.14)$$

is strictly less than one, strictly greater than one, or equals to one, respectively. In the subcritical and critical cases, the solution s of (4.13) equals to $\backslash(\mathbf{1}_-\backslash)$, while in the supercritical case $\backslash(s \leq \mathbf{1}_-, s \neq \mathbf{1}_-\backslash)$. Thus, the supercritical cases is the only interesting case. For the details, see (Hautphenne et al. 2008, 2011).

We apply the pure and relaxed Newton method for the example in Hautphenne et al. (2008). Let $\backslash(D_{\{0\}}\backslash)$, $\backslash(D_{\{1\}}\backslash)$, $\backslash(P_{\{0\}}\backslash)$, and $\backslash(\mathbf{d}\backslash)$ be defined by Figure-Fig. 3.

40314_2019_1021_Fig3_print.png

Fig. 3 Parameter for a family of unbalanced MBTs (Hautphenne et al. 2008)

The diagonal of $\backslash(D_{\{0\}}\backslash)$ is such that $\backslash(D_{\{0\}} \mathbf{1}_- + D_{\{1\}} \mathbf{1}_- + \mathbf{d} = 0\backslash)$. Define that $\backslash(a = -D_{\{0\}}^{-1} \mathbf{d}\backslash)$ and $\backslash(B = -D_{\{0\}}^{-1} R\backslash)$, where $\backslash(R_{\{i,9(j-1)+k\}} = [D_{\{1\}}]_{\{ii\}} [P_{\{1\}}]_{\{ij\}} [P_{\{0\}}]_{\{ik\}}\backslash)$ for $\backslash(i,j,k = 1,2,\dots,9\backslash)$. Then, the MBTE (4.13) is an order-convex equation on $\backslash(\mathbb{R}^n\backslash)$. Moreover, according to Hautphenne et al. (2008), (4.13) is supercritical if $\backslash(\delta\backslash)$ is greater than about 0.85. So, we experiment with $\backslash(\delta \geq 0.85\backslash)$.

Figure 4 shows the relative residuals of the pure and relaxed Newton iterations when $\backslash(\delta = 1.85\backslash)$ and $\backslash(\delta = 0.85 + 10^{-7}\backslash)$, respectively. Table 3 reports the numbers of two iterations for each $\backslash(\delta\backslash)$. From the experiments, we also see that the relaxed Newton iteration is less than the pure one. The efficiency on the iteration of the relaxed one is about 20%.

40314_2019_1021_Fig4_print.png

Fig. 4 The ~~convergences~~ convergence of the relaxed and pure Newton methods in Example 4.4

Table 3 The numbers of iterations for Example 4.4

$\backslash(\delta\backslash)$	Pure	Relaxed	$\backslash(\delta\backslash)$	Pure	Relaxed
$\backslash(0.85 + 10^{\{-0\}}\backslash)$	10	8	$\backslash(0.85 + 10^{\{-4\}}\backslash)$	21	17
$\backslash(0.85 + 10^{\{-1\}}\backslash)$	13	11	$\backslash(0.85 + 10^{\{-5\}}\backslash)$	23	18
$\backslash(0.85 + 10^{\{-2\}}\backslash)$	16	13	$\backslash(0.85 + 10^{\{-6\}}\backslash)$	23	19
$\backslash(0.85 + 10^{\{-3\}}\backslash)$	18	15	$\backslash(0.85 + 10^{\{-7\}}\backslash)$	23	19

Conclusion

For an order-convex matrix equation (2.4), it was proved that the convergence of the Newton sequence if the Fréchet derivative at the solution of F is nonsingular. We have proved the convergence even if the Fréchet derivative is singular. On the other hand, we have shown that the relaxed Newton method can be applied for not only matrix polynomial equations but also order-convex matrix equations. We also have shown that the relaxed Newton method for order-convex matrix equations is more efficient than the pure one on the iterations.

[40314_2019_1021_OnlinePDF.pdf](#)

Article Note

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP) (NRF-2017R1A5A1015722). The authors thank the anonymous referees for providing very useful suggestions for providing this paper.

References

- Alfa Attahiru Sule AS (2003) Combined elapsed time and matrix-analytic method for the discrete time $(GI/G/1)$ and $(GI^X/G/1)$ systems. *Queueing Syst* 45:5–25
- Bean Nigel GNG, Bright Leslie WLW, Latouche Guy G, Pearce Charles E MCEM, Pollett Philip KPK, Peter G (1997) Taylor. The quasi-stationary behavior of quasi-birth-and-death processes. *Ann Appl Probab* 7(1):134–155
- Benner Peter P, Byers Ralph R (1998) An exact line search method for solving generalized continuous-time algebraic Riccati equations. *IEEE Trans Automat Control* 43(1):101–107
- Dario A (2005) Bini DA, Guy Latouche G, and Beatrice Meini. Meini B (2005) Numerical methods for structured Markov chains. Oxford University Press, Oxford
- Geoffrey J Butler JB, Johnson Charles RCR, Wolkowicz Henry H (1985) Nonnegative solutions of a quadratic matrix equation arising from comparison theorems in ordinary differential equations. *SIAM J Algebraic Discrete Methods* 6(1):47–53
- Davis George J GJ (1981) Numerical solution of a quadratic matrix equation. *SIAM J Sci Comput* 2:164–175
- Davis George J GJ (1983) Algorithm 598: an algorithm to compute solvent of the matrix equation $(AX^2 + BX + C = 0)$. *ACM Trans Math Software Softw* 9(2):246–254
- Gohberg Israel, Lancaster Peter P, Rodman Leiba L (1982) Matrix Polynomials polynomials. Academic Press, Amsterdam
- Guo Chun C Hua H (2001) Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices. *SIAM J Matrix Anal Appl* 23(1):225–242
- Guo Chun C Hua H, Higham Nicholas J NJ (2007) Iterative solution of a nonsymmetric algebraic Riccati equation. *SIAM J Matrix Anal Appl* 29:396–412
- Guo Chun C Hua H, Laub Alan J AJ (2000) On the iterative solution of a class of nonsymmetric algebraic Riccati equations. *SIAM J Matrix Anal Appl* 22(2):376–391
- Guo Xiao X Xia X, Lin Wen W Wei W, Shu-Fang Xu X (2006) A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation. *Numer Math* 103:393–412
- Haccou Patsy P, Jagers Peter P, Vatutin Vladimir A VA (2005) Branching Processes processes:

Variation variation, Growth growth, and Extinction extinction of Populations populations. Cambridge University Press, Cambridge

Hautphenne Sophie S, Latouche Guy G, Remiche Marie M Ange A (2008) Newton's iteration for the extinction probability of a Markovian binary tree. *Linear Algebra Appl* 428:2791–2804

Hautphenne Sophie S, Latouche Guy G, Remiche Marie M Ange A (2011) Algorithmic approach to the extinction probability of branching processes. *Methodol Comput Appl Probab* 13(1):171–192

Hautphenne Sophie S, Leibnitz Kenji K, Remiche Marie M Ange A (2006) Extinction probability in Peer-to-Peer file diffusion. *ACM SIGMETRICS Perform Eval Rev* 34:3–4

He Qi Q Ming M, Neuts Marcel F MF (2001) On the convergence and limits of certain matrix sequences arising in quasi-birth-and-death Markov chains. *J Appl Probab* 38(2):519–541

Higham Nicholas J NJ (2008) *Functions of Matrices matrices: Theory theory and Computation computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA

Higham Nicholas J NJ, Kim Hyun H Min M (2000) Numerical analysis of a quadratic matrix equation. *IMA J Numer Anal* 20(4):499–519

Higham Nicholas J NJ, Kim Hyun H Min M (2001) Solving a quadratic matrix equation by Newton's method with exact line searches. *SIAM J Matrix Anal Appl* 23:303–316

Roger A Horn RA, Johnson CR (1985) *Horn and Charles R Matrix analysis*. Cambridge University Press, Johnson. Matrix Analysis Cambridge

Juang Jong J (1995) Existence of algebraic matrix Riccati equations arising in transport theory. *Linear Algebra Appl* 230:89–100

Kim Hyun H Min M (2008) Convergence of Newton's method for solving a class of quadratic matrix equations. *Honam Math J* 30(2):399–409

Kimmel Marek M, Axelrod David E DE (2002) *Branching Processes processes in Biology biology*. Springer-Verlag, New York

Lancaster Peter P (1966) *Lambda-matrices and Vibrating Systems vibrating systems*. Pergamon Press, Oxford

Lancaster Peter P, Tismenetsky Miron M (1985) *The Theory theory of Matrices matrices with Applications applications*, 2nd edn. Academic Press, Amsterdam

Latouche Guy L, Ramaswami Vaidyanathan R (1999) *Introduction to Matrix Analytic Methods in Stochastic Modeling matrix analytic methods in stochastic modeling*. ASA-SIAM

Ortega James M MJM, Rheinboldt Werner G WC (2000) *Iterative Solution solution of Nonlinear Equations in Several Variables nonlinear equations in several variables*. SIAM

Poole George G, Boullion Thomas T (1974) A survey on \mathbb{M} -matrices. *SIAM Rev* 16(4):419–427

Rogers LCG (1994) Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains. *Ann Appl Probab* 4(2):390–413, 05

Seo Jong Hyeon JH, Ki Hyun H Min M, m (2008) Solving matrix polynomials by Newton's method with exact line searches. *J Korean Soc Ind Appl Math* 12(2):55–68

Seo Jong J Hyeon H, Kim Hyun H Min M (2014) Convergence of pure and relaxed Newton methods for solving a matrix polynomial equation arising in stochastic models. *Linear Algebra Appl* 440:34–49

Seo Sang S Hyup H, Seo Jong J Hyun H, Kim Hyun H Min M (2013) Newton's method for solving a quadratic matrix equation with special coefficient matrices. *Honam Math J* 35(3):417–433

Vandergraft James S JS (1967) Newton's method for convex operators in partially ordered spaces. *SIAM J Numer Anal* 4(3):406–432

Xue Jungong J, Shufang Xu X, Li Ren R Gang C (2012) Accurate solutions of \mathbb{M} -matrix algebraic Riccati equations. *Numer Math* 120(4):671–700