Linear Algebra and its Applications 432 (2010) 283–291

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Convergence rates of some iterative methods for nonsymmetric algebraic Riccati equations arising in transport theory

Chun-Hua Guo [a,*,1], Wen-Wei Lin [b,2]

[a] Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
[b] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

A R T I C L E I N F O

Article history:
Received 19 May 2009
Accepted 6 August 2009
Available online 2 September 2009

Submitted by C.K. Li

AMS classification:
15A24
65F10

Keywords:
Nonsymmetric algebraic Riccati equations
Fixed-point iterations
Minimal positive solution
Convergence rate

A B S T R A C T

We determine and compare the convergence rates of various fixed-point iterations for finding the minimal positive solution of a class of nonsymmetric algebraic Riccati equations arising in transport theory.

1. Introduction

In transport theory, we encounter nonsymmetric algebraic Riccati equations of the form

$$XCX - XD - AX + B = 0 \tag{1}$$

(see [13]), where $A, B, C, D \in \mathbb{R}^{n\times n}$ are given by

$$A = \Delta - eq^T, \quad B = ee^T, \quad C = qq^T, \quad D = \Gamma - qe^T,$$

* Corresponding author.
E-mail addresses: chguo@math.uregina.ca (C.-H. Guo), wwlin@math.nctu.edu.tw (W.-W. Lin).

1 The work of this author was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada and by the National Center for Theoretical Sciences in Taiwan.
2 The work of this author was partially supported by the National Science Council and the National Center for Theoretical Sciences in Taiwan.

0024-3795/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2009.08.004

with $\Delta = \operatorname{diag}(\delta_1, \delta_2, \ldots, \delta_n)$, $\Gamma = \operatorname{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n)$, $e = (1, 1, \ldots, 1)^T$, $q = (q_1, q_2, \ldots, q_n)^T$. Here

$$q_i = \frac{c_i}{2w_i}, \tag{2}$$

with

$$0 < w_n < \cdots < w_2 < w_1 < 1, \quad c_i > 0 \quad (i = 1, 2, \ldots, n), \quad \sum_{i=1}^{n} c_i = 1,$$

and

$$\delta_i = \frac{1}{cw_i(1+\alpha)}, \quad \gamma_i = \frac{1}{cw_i(1-\alpha)}, \tag{3}$$

where $0 < c \leqslant 1$, $0 \leqslant \alpha < 1$. For descriptions on how these equations arise in transport theory, see [13] and references cited therein.

For any matrices $A, B \in \mathbb{R}^{m \times n}$, we write $A \geqslant B (A > B)$ if $a_{ij} \geqslant b_{ij} (a_{ij} > b_{ij})$ for all i, j. We can then define positive matrices, nonnegative matrices, etc. The existence of positive solutions of (1) has been shown in [12,13]. However, only the minimal positive solution is physically meaningful. More general nonsymmetric algebraic Riccati equations have been studied in [6,7,9,11]. In particular, the existence of positive solutions is proved for the wider class in [6,7] using elementary matrix theory.

Due to the special structures of the equation (1), its minimal positive solution can be found by iterative methods with $O(n^2)$ complexity each iteration, see [1,2,4,12,14]. The case $(\alpha, c) = (0, 1)$ is the most difficult to handle, and has been solved efficiently by using a shift technique in [4]. If $(\alpha, c) \neq (0, 1)$, the fixed-point iterations in [1,2,14] are linearly convergent. These methods are very simple and requires only $4n^2$ flops each iteration. The methods in [4,15] are more complicated. Those methods are quadratically convergent, but require more computations each iteration. Generally speaking, those methods should be used when (α, c) is relatively close to $(0, 1)$. Otherwise, the fixed-point iterations in [1,2,14] are usually adequate, and even more efficient. In this paper we further study the methods in [1,2,14]. We show that the NBGS method in [1] is the best one among these methods. In particular, we show that the NBGS method is twice as fast as the NBJ method in [1].

2. Preliminaries

It is shown in [14] that the minimal positive solution X^* of (1) has the form

$$X^* = T \circ (u^*(v^*)^T).$$

Here $\circ$ is the Hadamard product, $T = [t_{ij}]$ with $t_{ij} = 1/(\delta_i + \gamma_j)$, and (u^*, v^*) is the minimal positive solution of the vector equations

$$\begin{cases} u = u \circ (Pv) + e, \\ v = v \circ (Qu) + e, \end{cases} \tag{4}$$

where $P = [p_{ij}]$ and $Q = [q_{ij}]$ are $n \times n$ positive matrices given by

$$p_{ij} = \frac{q_j}{\delta_i + \gamma_j}, \quad q_{ij} = \frac{q_j}{\delta_j + \gamma_i}.$$

Four simple iterative methods have been proposed for finding the minimal solution (u^*, v^*). Each of them starts with $(u^{(0)}, v^{(0)}) = (0, 0)$. The simplest of them is the *simple iteration* (SI)

$$\begin{cases} u^{(k+1)} = u^{(k)} \circ \left(Pv^{(k)}\right) + e, \\ v^{(k+1)} = v^{(k)} \circ (Qu^{(k)}) + e. \end{cases} \tag{5}$$

It is shown in [14] that the sequence $\left\{\left(u^{(k)}, v^{(k)}\right)\right\}$ is strictly and monotonically increasing, and converges to (u^*, v^*). Later a *modified simple iteration* (MSI) is proposed in [2]:

$$\begin{cases} u^{(k+1)} = u^{(k)} \circ \left(Pv^{(k)}\right) + e, \\ v^{(k+1)} = v^{(k)} \circ \left(Qu^{(k+1)}\right) + e, \end{cases} \tag{6}$$

It is shown in [2] that the sequence $\left\{\left(u^{(k)}, v^{(k)}\right)\right\}$ is strictly and monotonically increasing, and converges to (u^*, v^*). Recently, two more methods are proposed in [1]. They are the *nonlinear block Jacobi* (NBJ) method

$$\begin{cases} u^{(k+1)} = u^{(k+1)} \circ \left(Pv^{(k)}\right) + e, \\ v^{(k+1)} = v^{(k+1)} \circ \left(Qu^{(k)}\right) + e, \end{cases} \tag{7}$$

and the *nonlinear block Gauss–Seidel* (NBGS) method

$$\begin{cases} u^{(k+1)} = u^{(k+1)} \circ \left(Pv^{(k)}\right) + e, \\ v^{(k+1)} = v^{(k+1)} \circ \left(Qu^{(k+1)}\right) + e. \end{cases} \tag{8}$$

It is shown in [1] that the sequence $\left\{\left(u^{(k)}, v^{(k)}\right)\right\}$ from either NBJ or NBGS is strictly and monotonically increasing, and converges to (u^*, v^*).

When there is a need to distinguish $\left(u^{(k)}, v^{(k)}\right)$ from SI, MSI, NBJ, or NBGS, they will be denoted by $\left(u_S^{(k)}, v_S^{(k)}\right)$, $\left(u_M^{(k)}, v_M^{(k)}\right)$, $\left(u_J^{(k)}, v_J^{(k)}\right)$, $\left(u_G^{(k)}, v_G^{(k)}\right)$, respectively.

The following result has been proved in [5].

Theorem 1. *For each* $k \geqslant 0$,

$$0 \leqslant u_S^{(k)} \leqslant u_J^{(k)} \leqslant u_G^{(k)}, \quad 0 \leqslant v_S^{(k)} \leqslant v_J^{(k)} \leqslant v_G^{(k)}.$$

It is easy to show that strict inequalities hold in Theorem 1 for $k \geqslant 2$. The next result is given in [2].

Theorem 2. *For each* $k \geqslant 0$,

$$u_S^{(k)} \leqslant u_M^{(k)}, \quad v_S^{(k)} \leqslant v_M^{(k)}.$$

Moreover, strict inequalities hold for $k \geqslant 3$.

It is easy to show by example that there is no similar comparison result for $\left(u_M^{(k)}, v_M^{(k)}\right)$ and $\left(u_J^{(k)}, v_J^{(k)}\right)$. However, it is easy to prove the following comparison result for $\left(u_M^{(k)}, v_M^{(k)}\right)$ and $\left(u_G^{(k)}, v_G^{(k)}\right)$.

Theorem 3. *For each* $k \geqslant 0$,

$$u_M^{(k)} \leqslant u_G^{(k)}, \quad v_M^{(k)} \leqslant v_G^{(k)}.$$

Moreover, strict inequalities hold for $k \geqslant 2$.

Proof. We have $u_M^{(0)} = u_G^{(0)} = 0$ and $v_M^{(0)} = v_G^{(0)} = 0$. It is easily seen that $u_M^{(1)} = u_G^{(1)} = e$ and $v_M^{(1)} = e$. By (4), $u^* = u^* \circ (Pv^*) + e$ and $v^* \circ (e - Qu^*) = e$. So $u^* > e$ and $0 < e - Qu^* < e - Qe < e$. It follows that $v_G^{(1)} > e$. Now assume $u_M^{(k)} \leqslant u_G^{(k)}$ and $v_M^{(k)} \leqslant v_G^{(k)}$ $(k \geqslant 1)$. Then

$$u_G^{(k+1)} = u_G^{(k+1)} \circ \left(Pv_G^{(k)}\right) + e > u_G^{(k)} \circ \left(Pv_G^{(k)}\right) + e \geqslant u_M^{(k)} \circ \left(Pv_M^{(k)}\right) + e = u_M^{(k+1)},$$

$$v_G^{(k+1)} = v_G^{(k+1)} \circ \left(Qu_G^{(k+1)}\right) + e > v_G^{(k)} \circ \left(Qu_G^{(k+1)}\right) + e \geqslant v_M^{(k)} \circ \left(Qu_M^{(k+1)}\right) + e = v_M^{(k+1)}.$$

We have thus proved the result by induction. □

Although strict inequalities hold in Theorems 1–3 after a few iterations, the asymptotic rates of convergence could still be the same for these methods. Thus a careful convergence rate analysis is needed.

3. Convergence rate analysis

Let

$$w^{(k)} = \begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix}, \quad w^* = \begin{bmatrix} u^* \\ v^* \end{bmatrix}.$$

Then each of the iterations (5)–(8) can be written as

$$w^{(k+1)} = \mathcal{F}(w^{(k)}),$$

where $\mathcal{F}$ is a mapping from $\mathbb{R}^{2n}$ into itself and w^* is a fixed point of $\mathcal{F}$. We let

$$d^{(k)} = w^* - w^{(k)},$$

and will find the matrix $L^{(k)}$ in the error relation

$$d^{(k+1)} = L^{(k)} d^{(k)} \tag{9}$$

for each of the four iterations. The Fréchet derivative of the mapping $\mathcal{F}$ at w^* will then be given by

$$\mathcal{F}'(w^*) = \lim_{k \to \infty} L^{(k)}.$$

The derivative will be denoted by $\mathcal{F}_S'(w^*)$, $\mathcal{F}_M'(w^*)$, $\mathcal{F}_J'(w^*)$, and $\mathcal{F}_G'(w^*)$ for SI, MSI, NBJ and NBGS, respectively.

For SI, we have

$$\begin{aligned} u^* - u^{(k+1)} &= (u^* \circ (Pv^*) + e) - \left(u^{(k)} \circ \left(Pv^{(k)}\right) + e\right) \\ &= \left(Pv^{(k)}\right) \circ \left(u^* - u^{(k)}\right) + u^* \circ \left(P\left(v^* - v^{(k)}\right)\right), \end{aligned} \tag{10}$$

and similarly

$$v^* - v^{(k+1)} = v^* \circ \left(Q\left(u^* - u^{(k)}\right)\right) + \left(Qu^{(k)}\right) \circ \left(v^* - v^{(k)}\right).$$

Thus (9) holds with

$$L^{(k)} = \begin{bmatrix} \operatorname{diag}\left(Pv^{(k)}\right) & \operatorname{diag}(u^*)\, P \\ \operatorname{diag}(v^*)\, Q & \operatorname{diag}\left(Qu^{(k)}\right) \end{bmatrix},$$

and we have

$$\mathcal{F}_S'(w^*) = \begin{bmatrix} \operatorname{diag}(Pv^*) & \operatorname{diag}(u^*)P \\ \operatorname{diag}(v^*)Q & \operatorname{diag}(Qu^*) \end{bmatrix}.$$

For MSI, the mapping $\mathcal{F}$ is given by

$$\mathcal{F}\begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix} = \begin{bmatrix} u^{(k)} \circ \left(Pv^{(k)}\right) + e \\ v^{(k)} \circ \left(Q\left(u^{(k)} \circ \left(Pv^{(k)}\right) + e\right)\right) + e \end{bmatrix}.$$

So the expression for $u^* - u^{(k+1)}$ is still given by (10). But we now have

$$\begin{aligned} v^* - v^{(k+1)} &= v^* \circ (Qu^*) - v^{(k)} \circ \left(Q\left(u^{(k)} \circ \left(Pv^{(k)}\right) + e\right)\right) \\ &= v^* \circ (Qu^*) - v^{(k)} \circ (Qu^*) \\ &\quad + v^{(k)} \circ (Q\,(u^* \circ (Pv^*) + e)) - v^{(k)} \circ \left(Q\left(u^{(k)} \circ \left(Pv^{(k)}\right) + e\right)\right) \\ &= (Qu^*) \circ \left(v^* - v^{(k)}\right) \\ &\quad + v^{(k)} \circ \left(Q\left(\left(Pv^{(k)}\right) \circ \left(u^* - u^{(k)}\right) + u^* \circ \left(P\left(v^* - v^{(k)}\right)\right)\right)\right). \end{aligned}$$

Thus (9) holds with

$$L^{(k)} = \begin{bmatrix} \operatorname{diag}\left(Pv^{(k)}\right) & \operatorname{diag}(u^*)P \\ \operatorname{diag}\left(v^{(k)}\right) Q \operatorname{diag}\left(Pv^{(k)}\right) & \operatorname{diag}(Qu^*) + \operatorname{diag}\left(v^{(k)}\right) Q \operatorname{diag}(u^*)P \end{bmatrix},$$

and we have

$$\mathcal{F}'_M(w^*) = \begin{bmatrix} \operatorname{diag}(Pv^*) & \operatorname{diag}(u^*)P \\ \operatorname{diag}(v^*)Q\operatorname{diag}(Pv^*) & \operatorname{diag}(Qu^*) + \operatorname{diag}(v^*)Q\operatorname{diag}(u^*)P \end{bmatrix}.$$

For NBJ, the mapping $\mathcal{F}$ is given by

$$\mathcal{F}\begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix} = \begin{bmatrix} e \Big/ \left(e - Pv^{(k)}\right) \\ e \Big/ \left(e - Qu^{(k)}\right) \end{bmatrix},$$

where / is componentwise division. It is easy to find that (9) holds with

$$L^{(k)} = \begin{bmatrix} 0 & \operatorname{diag}\left(u^* \circ u^{(k+1)}\right) P \\ \operatorname{diag}\left(v^* \circ v^{(k+1)}\right) Q & 0 \end{bmatrix}.$$

Thus

$$\mathcal{F}'_J(w^*) = \begin{bmatrix} 0 & \operatorname{diag}(u^* \circ u^*)P \\ \operatorname{diag}(v^* \circ v^*)Q & 0 \end{bmatrix}.$$

For NBGS, the mapping $\mathcal{F}$ is given by

$$\mathcal{F}\begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix} = \begin{bmatrix} e \Big/ \left(e - Pv^{(k)}\right) \\ e \Big/ \left(e - Q\left(e \Big/ \left(e - Pv^{(k)}\right)\right)\right) \end{bmatrix}.$$

We find that (9) holds with

$$L^{(k)} = \begin{bmatrix} 0 & \operatorname{diag}\left(u^* \circ u^{(k+1)}\right) P \\ 0 & \operatorname{diag}\left(v^* \circ v^{(k+1)}\right) Q \operatorname{diag}\left(u^* \circ u^{(k+1)}\right) P \end{bmatrix},$$

and that

$$\mathcal{F}'_G(w^*) = \begin{bmatrix} 0 & \operatorname{diag}(u^* \circ u^*)P \\ 0 & \operatorname{diag}(v^* \circ v^*)Q\operatorname{diag}(u^* \circ u^*)P \end{bmatrix}.$$

We now prove the following result about the rate of convergence.

Theorem 4. *For each of the iterations (5)–(8), we have*

$$\limsup_{k\to\infty} \sqrt[k]{\|d^{(k)}\|} = \rho(\mathcal{F}'(w^*)),$$

where $\|\cdot\|$ is any matrix norm and $\rho(\cdot)$ denotes the spectral radius.

Proof. For each iterative method we have for all $k \geqslant 0$

$$0 \leqslant L^{(k)} \leqslant L^{(k+1)} \leqslant \mathcal{F}'(w^*).$$

Thus

$$d^{(k)} = L^{(k-1)} \cdots L^{(1)} L^{(0)} d^{(0)} \leqslant (\mathcal{F}'(w^*))^k d^{(0)}.$$

So

$$\limsup_{k\to\infty} \sqrt[k]{\|d^{(k)}\|} \leqslant \limsup_{k\to\infty} \sqrt[k]{\|(\mathcal{F}'(w^*))^k\|} = \rho(\mathcal{F}'(w^*)).$$

Also, for any $k \geqslant l \geqslant 0$

$$d^{(k)} \geqslant \left(L^{(l)}\right)^{k-l} \left(L^{(0)}\right)^l d^{(0)}.$$

Note that $\left(L^{(0)}\right)^l d^{(0)} = \left(L^{(0)}\right)^l w^* > 0$. We can then prove that

$$\limsup_{k\to\infty} \sqrt[k]{\|d^{(k)}\|} \geqslant \rho(\mathcal{F}'(w^*)),$$

as in the proof of [10, Theorem 3.2]. □

The above convergence rate analysis reveals the following interesting result.

Theorem 5. *In terms of asymptotic rate of convergence, the NBGS method is twice as fast as the NBJ method.*

Proof. Note that

$$(\mathcal{F}_J'(w^*))^2 = \begin{bmatrix} \mathrm{diag}(u^* \circ u^*)P\mathrm{diag}(v^* \circ v^*)Q & 0 \\ 0 & \mathrm{diag}(v^* \circ v^*)Q\mathrm{diag}(u^* \circ u^*)P \end{bmatrix}.$$

So

$$(\rho(\mathcal{F}_J'(w^*)))^2 = \rho(\mathrm{diag}(v^* \circ v^*)Q\mathrm{diag}(u^* \circ u^*)P) = \rho(\mathcal{F}_G'(w^*)), \tag{11}$$

as required. □

Remark. The above theorem explains the numerical results for NBJ and NBGS presented in Tables 1 and 2 in [1], where the number of iterations required for NBGS is roughly half of that for NBJ.

The Riccati equation (1) contains two parameters c and α, $0 < c \leqslant 1$ and $0 \leqslant \alpha < 1$. We now examine the effect of these parameters on the rate of convergence, with $c_i, w_i (i = 1, \ldots, n)$ unchanged.

Theorem 6. *For each of the methods SI, MSI, NBJ, and NBGS, if c and α are changed such that $c(1+\alpha)$ and $c(1-\alpha)$ are decreasing with at least one of them strictly, then $\rho(\mathcal{F}'(w^*))$ is strictly decreasing.*

Proof. Under the assumption, the matrices P and Q are strictly decreasing. Using induction, we see easily from the SI method that u^* and v^* are also decreasing. We then see from (4) that at least one component of u^* or v^* is strictly decreasing. Note that $\mathcal{F}'(w^*)$ is an irreducible nonnegative matrix for SI, MSI, and NBJ, and that the (2, 2) block of $\mathcal{F}'(w^*)$ is an irreducible nonnegative matrix for NBGS. It follows from the Perron–Frobenius theory [3,16] that $\rho(\mathcal{F}'(w^*))$ is strictly decreasing. □

Remark. In Table 1 of [1] the number of iterations required for SI, NBJ, NBGS are reported for $(\alpha, c) = (10^{-6}, 1-10^{-6})$, $(0.001, 0.999)$, $(0.005, 0.995)$, $(0.1, 0.9)$, $(0.5, 0.5)$ (in this order). The results there show that the number of iterations decreases significantly for each method as (α, c) changes. This is explained (at least partially) by Theorem 6 since both $c(1+\alpha)$ and $c(1-\alpha)$ decrease significantly as (α, c) changes. Similarly, Theorem 6 explains the numerical results given in Tables 3.1 and 3.2 of [2] for SI and MSI, where (α, c) takes the values $(10^{-8}, 1-10^{-6})$, $(0.001, 0.999)$, $(0.01, 0.99)$, $(0.5, 0.5)$, $(0.85, 0.1)$.

Our main purpose in what follows is to show that NBGS is strictly faster than MSI (in terms of asymptotic rate of convergence) when $(\alpha, c) \neq (0, 1)$ and that the convergence of NBGS is still sublinear when $(\alpha, c) = (0, 1)$.

Let

$$K = I - \mathcal{F}_S'(w^*) = \begin{bmatrix} I - \mathrm{diag}(Pv^*) & -\mathrm{diag}(u^*)P \\ -\mathrm{diag}(v^*)Q & I - \mathrm{diag}(Qu^*) \end{bmatrix},$$

where I is an identity matrix of proper dimension. By definition, K is a nonsingular M-matrix if $\rho(\mathcal{F}_S'(w^*)) < 1$ and is a singular M-matrix if $\rho(\mathcal{F}_S'(w^*)) = 1$.

Lemma 7. *K is a nonsingular M-matrix if $(\alpha, c) \neq (0, 1)$, and is a singular M-matrix if $(\alpha, c) = (0, 1)$.*

Proof. The minimal positive solution X^* of (1) can be obtained by the fixed-point iteration

$$\Delta X_{k+1} + X_{k+1}\Gamma = X_k C X_k + B + eq^T X_k + X_k q e^T, \quad k = 0, 1, \ldots,$$

with $X_0 = 0$. Let the sequences $\{u^{(k)}\}$ and $\{v^{(k)}\}$ be obtained by (5). Then we have [14]

$$X_k = T \circ (u^{(k)} \left(v^{(k)}\right)^T), \quad u^{(k+1)} = X_k q + e, \quad v^{(k+1)} = X_k^T q + e.$$

It follows that X_k converges to X^* linearly if and only if $w^{(k)}$ converges to w^* linearly, which is the same as $\rho(\mathcal{F}_S'(w^*)) < 1$ by Theorem 4. On the other hand, by [10, Theorems 3.2 and 3.3] X_k converges to X^* linearly if and only if the matrix M_S in [10] is a nonsingular M-matrix. By [6, Propositions 3.4 and 4.9] and [8, Theorem 2.5], the matrix M_S in [10] is a nonsingular M-matrix if and only if $(\alpha, c) \neq (0, 1)$. We have thus proved that $\rho(\mathcal{F}_S'(w^*)) < 1$ when $(\alpha, c) \neq (0, 1)$ and $\rho(\mathcal{F}_S'(w^*)) = 1$ when $(\alpha, c) = (0, 1)$. □

We now consider four different regular splittings [16] of the matrix K: $K = M_i - N_i$, $i = 1, 2, 3, 4$, where

$$M_1 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \quad N_1 = \begin{bmatrix} \mathrm{diag}(Pv^*) & \mathrm{diag}(u^*)P \\ \mathrm{diag}(v^*)Q & \mathrm{diag}(Qu^*) \end{bmatrix},$$
$$M_2 = \begin{bmatrix} I & 0 \\ -\mathrm{diag}(v^*)Q & I \end{bmatrix}, \quad N_2 = \begin{bmatrix} \mathrm{diag}(Pv^*) & \mathrm{diag}(u^*)P \\ 0 & \mathrm{diag}(Qu^*) \end{bmatrix},$$
$$M_3 = \begin{bmatrix} I - \mathrm{diag}(Pv^*) & 0 \\ 0 & I - \mathrm{diag}(Qu^*) \end{bmatrix}, \quad N_3 = \begin{bmatrix} 0 & \mathrm{diag}(u^*)P \\ \mathrm{diag}(v^*)Q & 0 \end{bmatrix},$$
$$M_4 = \begin{bmatrix} I - \mathrm{diag}(Pv^*) & 0 \\ -\mathrm{diag}(v^*)Q & I - \mathrm{diag}(Qu^*) \end{bmatrix}, \quad N_4 = \begin{bmatrix} 0 & \mathrm{diag}(u^*)P \\ 0 & 0 \end{bmatrix}.$$

Lemma 8. $\mathcal{F}_S'(w^*) = M_1^{-1}N_1, \mathcal{F}_M'(w^*) = M_2^{-1}N_2, \mathcal{F}_J'(w^*) = M_3^{-1}N_3, \mathcal{F}_G'(w^*) = M_4^{-1}N_4$.

Proof. We prove the last equality. The others can be proved more easily. Using the formula

$$\begin{bmatrix} A & 0 \\ C & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ -B^{-1}CA^{-1} & B^{-1} \end{bmatrix},$$

and noting that by (4)

$$(I - \mathrm{diag}(Pv^*))^{-1} = \mathrm{diag}(u^*), \quad (I - \mathrm{diag}(Qu^*))^{-1} = \mathrm{diag}(v^*),$$

we obtain

$$M_4^{-1} = \begin{bmatrix} \mathrm{diag}(u^*) & 0 \\ \mathrm{diag}(v^* \circ v^*)Q\mathrm{diag}(u^*) & \mathrm{diag}(v^*) \end{bmatrix}.$$

A direct computation then gives $M_4^{-1}N_4 = \mathcal{F}_G'(w^*)$. □

Theorem 9. *If $(\alpha, c) = (0, 1)$, then*

$$\rho(\mathcal{F}_S'(w^*)) = \rho(\mathcal{F}_M'(w^*)) = \rho(\mathcal{F}_J'(w^*)) = \rho(\mathcal{F}_G'(w^*)) = 1. \quad (12)$$

If $(\alpha, c) \neq (0, 1)$, *then*

$$\rho(\mathcal{F}_G'(w^*)) < \rho(\mathcal{F}_M'(w^*)) < \rho(\mathcal{F}_S'(w^*)) < 1,$$
$$\rho(\mathcal{F}_G'(w^*)) < \rho(\mathcal{F}_J'(w^*)) < \rho(\mathcal{F}_S'(w^*)) < 1.$$

Proof. Recall that the Fréchet derivatives are all nonnegative matrices. In view of Lemmas 7 and 8, we have as in the proof of [10, Theorem 3.3] that (12) holds if $(\alpha, c) = (0, 1)$ and that

$$\rho(\mathcal{F}_G'(w^*)) \leqslant \rho(\mathcal{F}_M'(w^*)) \leqslant \rho(\mathcal{F}_S'(w^*)) < 1,$$
$$\rho(\mathcal{F}_G'(w^*)) \leqslant \rho(\mathcal{F}_J'(w^*)) \leqslant \rho(\mathcal{F}_S'(w^*)) < 1$$

if $(\alpha, c) \neq (0, 1)$. When $(\alpha, c) \neq (0, 1)$, by the theory of nonnegative matrices we know that [16, Theorem 3.29]

$$\rho(M_i^{-1}N_i) = \frac{\rho(K^{-1}N_i)}{1 + \rho(K^{-1}N_i)}. \tag{13}$$

Since $0 \leqslant K^{-1}N_4 \leqslant K^{-1}N_2$, $K^{-1}N_2 > 0$, and $K^{-1}N_4 \neq K^{-1}N_2$, we have $\rho(K^{-1}N_4) < \rho(K^{-1}N_2)$ by the Perron–Frobenius theory. So $\rho(M_4^{-1}N_4) < \rho(M_2^{-1}N_2)$ by (13), which is the same as $\rho(\mathcal{F}_G'(w^*)) < \rho(\mathcal{F}_M'(w^*))$. Similarly, we can prove $\rho(\mathcal{F}_M'(w^*)) < \rho(\mathcal{F}_S'(w^*))$ and $\rho(\mathcal{F}_G'(w^*)) < \rho(\mathcal{F}_J'(w^*)) < \rho(\mathcal{F}_S'(w^*))$. Note that $\rho(\mathcal{F}_G'(w^*)) < \rho(\mathcal{F}_J'(w^*))$ also follows from (11) directly. □

4. Conclusion

In this paper we have further studied four fixed-point iterations for finding the minimal positive solution of the equation (1), which involves a pair of parameters (α, c) with $0 \leqslant \alpha < 1$ and $0 < c \leqslant 1$. These methods are all easy to use, and have the same low complexity each iteration. We have shown that the NBGS method in [1] is faster than the other three in terms of asymptotic rate of convergence when $(\alpha, c) \neq (0, 1)$. Existing results and a new result in this paper together show that the NBGS method also provides better approximation after every iteration. We have also shown that the convergence of the NBGS method is still sublinear when $(\alpha, c) = (0, 1)$. So one should use the methods in [4,15] when (α, c) is close to $(0, 1)$, and use the NBGS method otherwise.

References

[1] Z.-Z. Bai, Y.-H. Gao, L.-Z. Lu, Fast iterative schemes for nonsymmetric algebraic Riccati equations arising from transport theory, SIAM J. Sci. Comput. 30 (2008) 804–818.
[2] L. Bao, Y. Lin, Y. Wei, A modified simple iterative method for nonsymmetric algebraic Riccati equations arising in transport theory, Appl. Math. Comput. 181 (2006) 1499–1504.
[3] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.
[4] D.A. Bini, B. Iannazzo, F. Poloni, A fast Newton's method for a nonsymmetric algebraic Riccati equation, SIAM J. Matrix Anal. Appl. 30 (2008) 276–290.
[5] Y.-H. Gao, Theories and Algorithms for Several Quadratic Matrix Equations, Ph.D. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, 2007.
[6] C.-H. Guo, Nonsymmetric algebraic Riccati equations and Wiener–Hopf factorization for M-matrices, SIAM J. Matrix Anal. Appl. 23 (2001) 225–242.
[7] C.-H. Guo, A note on the minimal nonnegative solution of a nonsymmetric algebraic Riccati equation, Linear Algebra Appl. 357 (2002) 299–302.
[8] C.-H. Guo, On a quadratic matrix equation associated with an M-matrix, IMA J. Numer. Anal. 23 (2003) 11–27.
[9] C.-H. Guo, N.J. Higham, Iterative solution of a nonsymmetric algebraic Riccati equation, SIAM J. Matrix Anal. Appl. 29 (2007) 396–412.
[10] C.-H. Guo, A.J. Laub, On the iterative solution of a class of nonsymmetric algebraic Riccati equations, SIAM J. Matrix Anal. Appl. 22 (2000) 376–391.
[11] X.-X. Guo, W.-W. Lin, S.-F. Xu, A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation, Numer. Math. 103 (2006) 393–412.
[12] J. Juang, Existence of algebraic matrix Riccati equations arising in transport theory, Linear Algebra Appl. 230 (1995) 89–100.

[13] J. Juang, W.-W. Lin, Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices, SIAM J. Matrix Anal. Appl. 20 (1998) 228–243.
[14] L.-Z. Lu, Solution form and simple iteration of a nonsymmetric algebraic Riccati equation arising in transport theory, SIAM J. Matrix Anal. Appl. 26 (2005) 679–685.
[15] V. Mehrmann, H. Xu, Explicit solutions for a Riccati equation from transport theory, SIAM J. Matrix Anal. Appl. 30 (2008) 1339–1357.
[16] R.S. Varga, Matrix Iterative Analysis, second ed., Springer, Berlin, 2000.