Newton-Shamanskii Method for a Quadratic Matrix Equation Arising in Quasi-Birth-Death Problems

Pei-Chang Guo*

School of Mathematical Sciences, Peking University, Beijing 100871, China. Received 4 September 2014; Accepted (in revised version) 30 October 2014 Available online 5 November 2014

> **Abstract.** In order to determine the stationary distribution for discrete time quasi-birthdeath Markov chains, it is necessary to find the minimal nonnegative solution of a quadratic matrix equation. The Newton-Shamanskii method is applied to solve this equation, and the sequence of matrices produced is monotonically increasing and converges to its minimal nonnegative solution. Numerical results illustrate the effectiveness of this procedure.

AMS subject classifications: 65F30, 65H10

Key words: Quadratic matrix equation, quasi-birth-death problems, Newton-Shamanskii method, minimal nonnegative solution.

1. Introduction

Some necessary notation for this article is as follows. For any matrix $B = [b_{ij}] \in \mathbb{R}^{n \times n}$, $B \ge 0$ (B > 0) if $b_{ij} \ge 0$ ($b_{ij} > 0$) for all i, j; for any matrices $A, B \in \mathbb{R}^{n \times n}$, $A \ge B$ (A > B) if $a_{ij} \ge b_{ij}$ ($a_{ij} > b_{ij}$) for all i, j; for any vectors $x, y \in \mathbb{R}^n$, $x \ge y$ (x > y) if $x_i \ge y_i$ ($x_i > y_i$) for all $i = 1, \dots, n$; the vector with all entries one is denoted by $e - i.e. \ e = (1, 1, \dots, 1)^T$; and the identity matrix is denoted by I. The quadratic matrix equation (QME)

$$\mathcal{Q}(X) = AX^2 + BX + C = 0 \tag{1.1}$$

is considered, where $A, B, C, X \in \mathbb{R}^{n \times n}$, $A, B + I, C \ge 0$, A + B + I + C is irreducible and (A + B + C)e = e. This quadratic matrix equation arises in quasi-birth-death processes (QBD) [4], and its element-wise minimal nonnegative solution *S* is of particular interest. The rate ρ of a QBD Markov chain is defined by

$$\rho = p^T (B + I + 2A)e , \qquad (1.2)$$

where *p* is the stationary probability vector of the stochastic matrix A + B + I + C — i.e. $p^{T}(A + B + I + C) = p^{T}$ and $p^{T}e = 1$ (cf. the monograph [4] for further details). A QBD is

^{*}Corresponding author. Email address: gpeichang@126.com (Pei-Chang Guo)

http://www.global-sci.org/eajam

said to be positive recurrent if $\rho < 1$, null recurrent if $\rho = 1$ and transient if $\rho > 1$ — and throughout this article the QBD is assumed to be positive recurrent.

There have been several numerical methods proposed to solve the QME (1.1). Some linearly convergent fixed point iterations are summarised and analysed in Ref. [1] and references therein. Latouche [2] showed the application of Newton's algorithm is well defined, and that the matrix sequence is monotonically increasing and quadratically convergent. An invariant subspace method approximates the minimal nonnegative solution S quadratically, through approximating the left invariant subspace of a block companion matrix [4,9]. Latouche & Ramaswami [5] proposed a logarithmic reduction algorithm generating sequences of approximations that converge quadratically to S, based on a divide-and conquer strategy. Bini et al. [6–10] proposed a quadratically convergent and numerically stable algorithm for the computation of S based on a functional representation of cyclic reduction, which applies to general M/G/1-type Markov chains [16] and generalises the method of Ref. [5]. Poloni [12] studied several quadratic vector and matrix equations with nonnegativity hypotheses in a unified fashion, giving further insight into the equations. The Newton-Shamanskii method has been proposed for other equations — e.g. the vector equation arising in transport theory [13], the algebraic Riccati equation with four coefficient matrices forming a nonsingular *M*-matrix or an irreducible singular *M*-matrix [14], and the vector equation arising in Markovian binary trees [15].

In this article, the Newton-Shamanskii method is applied to the QME (1.1). Newton's method is recalled and the Newton-Shamanskii iterative procedure is presented in Section 2. Then in Section 3 it is shown that, starting with a suitable initial guess, the sequence of iterative matrices generated by the Newton-Shamanskii method is monotonically increasing and converges to the minimal nonnegative solution of the QME (1.1). Numerical results in Section 4 show that the Newton-Shamanskii method can be more efficient than the Newton method. Final conclusions are presented in Section 5.

2. Newton-Shamanskii Method

The function \mathscr{Q} in the QME (1.1) is a mapping from $\mathbb{R}^{n \times n}$ into itself, and the Fréchet derivative of \mathscr{Q} at *X* is a linear map $\mathscr{Q}'_X : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ given by

$$\mathscr{Q}_{X}^{'}(Z) = AZX + AXZ + BZ. \tag{2.1}$$

The second derivative $\mathscr{Q}_X'': \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ at *X* is given by

$$\mathscr{Q}_{X}^{''}(Z_{1}, Z_{2}) = AZ_{1}Z_{2} + AZ_{2}Z_{1}.$$
(2.2)

For given X_0 , the Newton sequence for the solution of $\mathcal{Q}(X) = 0$ is

$$X_{k+1} = X_k - (\mathscr{Q}'_{X_k})^{-1} \mathscr{Q}(X_k), \quad k = 0, 1, 2, \cdots,$$
(2.3)

provided that $\mathcal{Q}_{X_k}^{'}$ is invertible for all k. From Eq. (2.1), the Newton iteration (2.3) is equivalent to

$$\begin{cases} AZX_k + (AX_k + B)Z = -\mathcal{Q}(X_k), \\ X_{k+1} = X_k + Z, \quad k = 0, 1, 2, \cdots \end{cases}$$
(2.4)

or

388

$$AX_{k+1}X_k + (AX_k + B)X_{k+1} = AX_k^2 - C$$
, $k = 0, 1, 2, \cdots$. (2.5)

For the nonlinear equation $\mathscr{Q}(X) = 0$ with the minimal nonnegative solution *S*, the sequence generated by the Newton method converges quadratically and globally to the solution *S* [2]. However, the Newton method has a disadvantage — viz. at every iteration step, a Sylvester-type equation

$$A_1 X B_1^T + A_2 X B_2^T = E$$

must be solved. When the QZ algorithm is involved, the Bartels-Stewart or Hessenberg-Schur methods can be employed to solve this equation [3] — i.e. a transformation method that employs the QZ algorithm may be invoked, to structure the equation in such a way that it can be solved column-wise by a back substitution technique. However, the work count of floating point operations involved in the QZ algorithm is large compared with the back substitution [3], and reusing the special coefficient matrix structure form produced by QZ algorithm is more efficient. The preferred Newton-Shamanskii algorithm for the QME (1.1) is as follows.

Newton-Shamanskii algorithm for the QME (1.1)

Given an initial value X_0 , for $k = 0, 1, \cdots$

$$X_{k,0} = X_k - (\mathscr{Q}'_{X_k})^{-1} \mathscr{Q}(X_k), \qquad (2.6)$$

$$X_{k,s} = X_{k,s-1} - (\mathscr{Q}_{X_k})^{-1} \mathscr{Q}(X_{k,s-1}), \quad s = 1, 2, \cdots, n_k,$$
(2.7)

$$X_{k+1} = X_{k,n_k} \,. \tag{2.8}$$

3. Convergence Analysis

There is monotone convergence when the Newton-Shamanskii method is applied to the QME (1.1).

3.1. Preliminary

Let us first recall that a real square matrix *A* is a *Z*-matrix if all its off-diagonal elements are nonpositive, and can be written as sI - B with $B \ge 0$. Moreover, a *Z*-matrix *A* is called an *M*-matrix if $s \ge \rho(B)$, where $\rho(\cdot)$ is the spectral radius; it is a singular *M*-matrix if $s = \rho(B)$, and a nonsingular *M*-matrix if $s > \rho(B)$. The following result from Ref. [17] is to be exploited.

Lemma 3.1. For a Z-matrix A, the following statements are equivalent:

(a) A is a nonsingular M-matrix;

(b) $A^{-1} \ge 0$;

Newton-Shamanskii Method for a Quadratic Matrix Equation Arising in Quasi-Birth-Death Problems 389

- (c) Av > 0 for some vector v > 0;
- (d) All eigenvalues of A have positive real parts.

The following result is also well known [17].

Lemma 3.2. Let A be a nonsingular M-matrix. If $B \ge A$ is a Z-matrix, then B is a nonsingular M-matrix. Moreover, $B^{-1} \le A^{-1}$.

The minimal nonnegative solution *S* for the QME (1.1) may also be recalled — cf. Refs. [2,4] for details.

Theorem 3.1. If the quasi-birth-death process is positive recurrent (i.e. if the rate ρ defined by Eq. (1.2) is less than 1), then the matrix

$$-[(S^T \otimes A + I \otimes AS) + I \otimes B]$$

is a nonsingular M-matrix.

3.2. Monotone convergence

The following lemma displays the monotone convergence properties of the Newton iteration for the QME (1.1).

Lemma 3.3. Consider a matrix X such that

- (i) $\mathscr{Q}(X) \geq 0$,
- (ii) $0 \le X \le S$,
- (iii) $-[(X^T \otimes A + I \otimes AX) + I \otimes B]$ is a nonsingular M-matrix.

Then the matrix

$$Y = X - (\mathcal{Q}_{X}^{'})^{-1} \mathcal{Q}(X)$$
(3.1)

exists such that

- (a) $\mathscr{Q}(Y) \geq 0$,
- (b) $0 \le X \le Y \le S$,
- (c) $-[(Y^T \otimes A + I \otimes AY) + I \otimes B]$ is a nonsingular M-matrix.

Proof. \mathscr{Q}'_X is invertible and the matrix *Y* is well defined, from (iii) and Lemma 3.1. Since $\mathscr{Q}(X) \ge 0$ and $-[(X^T \otimes A + I \otimes AX) + I \otimes B]^{-1} \ge 0$ from (iii) and Lemma 3.1, it follows that $vec(Y) \ge vec(X)$ and thus $Y \ge X$. From Eq. (3.1) and the Taylor formula,

$$\mathcal{Q}(Y) = \mathcal{Q}(X) + \mathcal{Q}'_X(Y - X) + \frac{1}{2}\mathcal{Q}''_X(Y - X, Y - X)$$
$$= \frac{1}{2}\mathcal{Q}''_X(Y - X, Y - X)$$
$$= \mathcal{A}(Y - X)^2 \ge 0,$$

P.-C. Guo

and (b) may be proven as follows. From

$$AYX + (AX + B)Y = AX^2 - C ,$$

equivalent to Eq. (3.1), and the equation

$$AS^2 + BS + C = 0,$$

it follows that

$$A(Y-S)X + (AX+B)(Y-S) = AX^2 - C - ASX - AXS - BS$$
$$= A(X-S)(X-S)$$
$$> 0.$$

It is notable that $-[(X^T \otimes A + I \otimes AX) + I \otimes B]$ is a nonsingular *M*-matrix, so $vec(S - Y) \ge 0$ from Lemma 3.1 — i.e. $S - Y \ge 0$. Now $Y \ge X$, so (b) follows. From $0 \le Y \le S$,

$$-[(Y^T \otimes A + I \otimes AY) + I \otimes B] \ge -[(S^T \otimes A + I \otimes AS) + I \otimes B],$$

and $-[(S^T \otimes A + I \otimes AS) + I \otimes B]$ is a nonsingular *M*-matrix. Consequently from Lemma 3.2, $-[(Y^T \otimes A + I \otimes AY) + I \otimes B]$ is a nonsingular *M*-matrix.

An extension of Lemma 3.3 provides the theoretical basis for the monotone convergence of the Newton-Shamanskii method for the QME (1.1).

Lemma 3.4. Consider a matrix X such that

- (i) $\mathscr{Q}(X) \geq 0$,
- (ii) $0 \le X \le S$,
- (iii) $-[(X^T \otimes A + I \otimes AX) + I \otimes B]$ is a nonsingular M-matrix.

Then for any matrix N where $0 \le N \le X$, the matrix

$$Y = X - (\mathscr{Q}_N)^{-1} \mathscr{Q}(X)$$
(3.2)

exists such that

- (a) $\mathscr{Q}(Y) \geq 0$,
- (b) $0 \le X \le Y \le S$,
- (c) $-[(Y^T \otimes A + I \otimes AY) + I \otimes B]$ is a nonsingular M-matrix.

Proof. Since $0 \le N \le X$,

$$-[(N^T \otimes A + I \otimes AN) + I \otimes B] \ge -[(X^T \otimes A + I \otimes AX) + I \otimes B].$$

From (iii) and Lemma 3.2, $\mathscr{Q}_{N}^{'}$ is invertible and the matrix *Y* is well defined such that $0 \le X \le Y$. Let

$$\hat{Y} = X - (\mathscr{Q}_{X})^{-1} \mathscr{Q}(X),$$

390

Newton-Shamanskii Method for a Quadratic Matrix Equation Arising in Quasi-Birth-Death Problems 391

such that $\hat{Y} \ge Y$ from Lemma 3.2. As also $\hat{Y} \le S$ from Lemma 3.3, (b) follows. Now

$$-[(\hat{Y}^T \otimes A + I \otimes A\hat{Y}) + I \otimes B]$$

is a nonsingular *M*-matrix from Lemma 3.3 and $\hat{Y} \ge Y$, and $-[(Y^T \otimes A + I \otimes AY) + I \otimes B]$ is a nonsingular *M*-matrix from Lemma 3.2. From the Taylor formula and also noting that $\mathcal{Q}_N'(Y-X) + \mathcal{Q}(X) = 0$,

$$\begin{aligned} \mathscr{Q}(Y) &= \mathscr{Q}(X) + \mathscr{Q}'_{X}(Y - X) + \frac{1}{2}\mathscr{Q}''_{X}(Y - X, Y - X) \\ &= \mathscr{Q}(X) + \mathscr{Q}'_{N}(Y - X) + (\mathscr{Q}'_{X} - \mathscr{Q}'_{N})(Y - X) + \frac{1}{2}\mathscr{Q}''_{X}(Y - X, Y - X) \\ &= (\mathscr{Q}'_{X} - \mathscr{Q}'_{N})(Y - X) + \frac{1}{2}\mathscr{Q}''_{X}(Y - X, Y - X) \\ &= \mathscr{Q}''_{X}(X - N, Y - X) + \frac{1}{2}\mathscr{Q}''_{X}(Y - X, Y - X) \\ &= A(X - N)(Y - X) + A(Y - X)(X - N) + A(Y - X)^{2} \\ &\geq 0. \end{aligned}$$

The monotone convergence result for the Newton-Shamanskii method applied to the QME (1.1) follows.

Theorem 3.2. Suppose that a matrix X_0 is such that

- (i) $\mathscr{Q}(X_0) \geq 0$,
- (ii) $0 \leq X_0 \leq S$,
- (iii) $-[(X_0^T \otimes A + I \otimes AX_0) + I \otimes B]$ is a nonsingular *M*-matrix.

Then the Newton-Shamanskii algorithm (2.6)–(2.8) generates a sequence $\{X_k\}$ such that $X_k \leq X_{k+1} \leq S$ for all $k \geq 0$, and $\lim_{k\to\infty} X_k = S$.

Proof. The proof is by mathematical induction. From Lemma 3.4,

$$X_0 \leq X_{0,0} \leq \cdots \leq X_{0,n_0} = X_1 \leq S ,$$

$$\mathcal{Q}(X_1) \geq 0 ,$$

and

$$-[(X_1^T \otimes A + I \otimes AX_1) + I \otimes B]$$

is a nonsingular M-matrix. Assuming

$$\mathscr{Q}(X_i) \ge 0$$
,
 $X_0 \le X_{0,0} \le \dots \le X_{0,n_0} = X_1 \le \dots \le X_{i-1,n_{i-1}} = X_i \le S$,

and that $-[(X_i^T \otimes A + I \otimes AX_i) + I \otimes B]$ is a nonsingular *M*-matrix, from Lemma 3.4

$$\mathcal{Q}(X_{i+1}) \ge 0,$$

$$X_i \le X_{i,0} \le \dots \le X_{i,n_i} = X_{i+1} \le S,$$

and $-[(X_{i+1}^T \otimes A + I \otimes AX_{i+1}) + I \otimes B]$ is a nonsingular *M*-matrix. By induction, the sequence $\{X_k\}$ is therefore monotonically increasing and bounded above by *S*, and so has a limit X_* such that $X_* \leq S$. Letting $i \to \infty$ in $X_{i+1} \geq X_{i,0} = X_i - (\mathscr{D}'_{X_i})^{-1} \mathscr{D}(X_i) \geq 0$, it follows that $\mathscr{D}(X_*) = 0$. Consequently, $X_* = S$ since $X_* \leq S$ and *S* is the minimal nonnegative solution of the QME (1.1).

4. Numerical Results

The Newton-Shamanskii method differs from Newton's method as the Fréchet derivative is not updated at each iteration. The coefficient matrix pairs of the Sylvester-type equation may be evaluated and reduced via the QZ algorithm after several inner iteration steps, so although more iterations are needed than for Newton's method the overall computational cost of the Newton-Shamanskii method is less. An example from Refs. [5, 10] was used to test the efficiency of the Newton-Shamskii method for the QME (1.1).

Example 4.1. A quasi-birth-death problem is defined by the $n \times n$ matrices A = W, B = W - I, $C = W + \delta I$, where *I* is the identity matrix, *W* is a matrix having null diagonal entries and constant off-diagonal entries and $0 < \delta < 1$. As observed in Ref. [5], the rate $\rho = p^T (B + I + 2A)e$ where $p^T (A + B + I + C) = p^T$ and $p^T e = 1$ is exactly $1 - \delta$.

As reported in Ref. [3], the Hesseberg-Schur method is faster than the Bartels-Stewart method when solving the general Sylvester-type equation

$$A_1 X B_1^T + A_2 X B_2^T = E ,$$

so the Hesseberg-Schur method is often adopted in Newton iteration. However, the Bartels-Stewart method was used in the Newton-Shamanskii calculations reported here, because the reduced coefficient matrix in the back substitution step may be reused. (In the first call to the QZ algorithm, A_1 was reduced to quasi-upper-triangular form.) The optimal scalars n_i in the Newton-like algorithm (2.7) were chosen without the benefit of any theoretical results, and the Fréchet derivative was updated every m = 2 steps. The number of evaluations of the Fréchet derivative in the algorithm corresponded to the outer iteration steps k + 1 in the Newton-Shamanskii algorithm for the approximate solution $x_{k,l}$.

Measures of the feasibility and effectiveness of the new method are the number of outer iteration steps (denoted by "it"), the elapsed CPU time in seconds ("time") and the normalised residual

NRes =
$$\frac{\|A\tilde{X}^2 + B\tilde{X} + C\|}{\|\tilde{X}\| (\|A\|\| \|\tilde{X}\| + \|B\|) + \|C\|},$$

where $\|\cdot\|$ denotes the infinity-norm of the matrix and \tilde{X} is an approximate solution to the minimal nonnegative solution of the QME (1.1). The numerical tests were performed on a laptop (2.4 Ghz and 2G Memory) with MATLAB R2013a. The initial value was $X_0 = 0$ and the stopping criterion was

$$||A\tilde{X}^2 + B\tilde{X} + C|| < 1e - 12$$

Three different δ values and three problem sizes (*n* values) were considered. Tables 1, 2 and 3 report the results obtained with sizes n = 20, n = 100 and n = 200, respectively. The results show that the Newton-Shamanskii method is more efficient than the Newton method.

δ	Method	time	it	NRes
5.0e-1	Newton	0.013	5	4.77e-16
5.0e-1	Newton-Shamanskii	0.009	3	2.38e-14
1.0e-1	Newton	0.031	7	1.61e-16
1.0e-1	Newton-Shamanskii	0.026	5	9.25e-16
1.0e-3	Newton	0.043	13	8.70e-16
1.0e-3	Newton-Shamanskii	0.036	9	3.00e-16

Table 1: Comparison of the numerical results when n = 20.

Table 2: Comparison of the numerical results when n = 100.

δ	Method	time	it	NRes
5.0e-1	Newton	0.142	5	1.24e-15
5.0e-1	Newton-Shamanskii	0.110	3	2.50e-14
1.0e-1	Newton	0.190	7	1.21e-15
1.0e-1	Newton-Shamanskii	0.168	5	1.60e-15
1.0e-3	Newton	0.444	13	1.60e-15
1.0e-3	Newton-Shamanskii	0.359	9	6.14e-16

Table 3: Comparison of the numerical results when n = 200.

δ	Method	time	it	NRes
5.0e-1	Newton	1.026	5	9.40e-15
5.0e-1	Newton-Shamanskii	0.746	3	2.34e-14
1.0e-1	Newton	1.433	7	2.18e-15
1.0e-1	Newton-Shamanskii	1.200	5	1.25e-15
1.0e-3	Newton	4.798	13	5.64e-15
1.0e-3	Newton-Shamanskii	4.271	9	2.50e-15

5. Conclusions

In this article, the application of the Newton-Shamanskii method to the quadratic matrix equation arising from the analysis of quasi-birth-death processes has been considerd. The convergence analysis shows that this method is feasible and the minimal nonnegative solution of the quadratic matrix equation can be obtained. Numerical calculations show that the Newton-Shamanskii method can outperform Newton's method.

Acknowledgments

The author would like to thank Prof. Shu-Fang Xu and Prof. Yun-Feng Cai for inspiring discussions and the two anonymous referees for carefully checking the paper and providing helpful suggestions. The author also wishes to thank Prof. Roger J. Hosking, the Co-Editor of EAJAM, for helping edit the file to present the final version.

References

- [1] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, SIAM, Philadelphia (1999).
- G. Latouche, Newton's iteration for non-linear equations in Markov chains, IMA J. Numer. Anal. 14 583-598 (1994).
- [3] J.D. Gardiner, A.J. Laub, J.J. Amato and C.B. Moler, *Solution of the Sylvester matrix equation* $AXB^T + CXD^T = E$, ACM Trans. Math. Software **18**, 223-231 (1992).
- [4] D.A. Bini, G. Latouche and B. Meini, *Numerical methods for structured Markov chains*, Oxford University Press, New York (2005).
- [5] G. Latouche and V. Ramaswami, *A logarithmic reduction algorithm for quasi-birth-death processes*, J. Appl. Prob. **30**, 650-674 (1993).
- [6] D.A. Bini and B. Meini, On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems, in Computations with Markov Chains, Kluwer, Dordrecht (1995).
- [7] D.A. Bini and B. Meini, On the solution of a nonlinear matrix equation arising in queueingproblems, SIAM J. Matrix Anal. Appl. 17, 906-926 (1996).
- [8] D.A. Bini and B. Meini, Improved cyclic reduction for solving queueing problems, Numer. Algorithms 15, 57-74 (1997).
- [9] B. Meini, Solving QBD problems: The cyclic reduction algorithm versus the invariant subspace method, Adv. Perf. Anal. 1, 215-225 (1998).
- [10] C. He, B. Meini and N.H. Rhee, A shifted cyclic reduction algorithm for quasi-birth-death problems, SIAM J. Matrix Anal. Appl. 23, 673-691 (2001).
- [11] D.A. Bini, G. Latouche and B. Meini, *Solving matrix polynomial equations arising in queueing problems*, Linear Algebra Appl. **340**, 225-244 (2002).
- [12] F. Poloni, *Quadratic vector equations*, Linear Algebra Appl. **438**,1627-1644 (2013).
- [13] Y. Lin and L. Bao, *Convergence analysis of the Newton-Shamanskii method for a nonsymmetric algebraic Riccati equation*, Numer. Linear Algebra Appl. **15**, 535-546 (2008).
- [14] C.-H. Guo, Monotone convergence of Newton-like methods for M-matrix algebraic Riccati equations, Numer. Algorithms 64, 295-309 (2013).
- [15] P.-C. Guo and S.-F. Xu, *The modified Newton-Shamanskii method for the solution of a quadratic vector equation arising in Markovian binary trees*, Calcolo, to appear.

Newton-Shamanskii Method for a Quadratic Matrix Equation Arising in Quasi-Birth-Death Problems 395

- [16] M.F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Dekker, New York (1989).
- [17] R. Varga, Matrix Iterative Analysis, Prentice-Hall (1962).