Article Submitted to IMA Journal of Numerical Analysis

On a quadratic matrix equation associated with an M-matrix*

Chun-Hua Guo

Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada (chguo@math.uregina.ca)

Abstract

We study the quadratic matrix equation $X^2 - EX - F = 0$, where E is diagonal and F is an M-matrix. Quadratic matrix equations of this type arise in noisy Wiener–Hopf problems for Markov chains. The solution of practical interest is a particular M-matrix solution. The existence and uniqueness of M-matrix solutions and numerical methods for finding the desired M-matrix solution are discussed by transforming the equation into an equation that belongs to a special class of nonsymmetric algebraic Riccati equations (AREs). We also discuss the general nonsymmetric ARE and describe how we can use the Schur method to find the stabilizing or almost stabilizing solution if it exists. The desired M-matrix solution of the quadratic matrix equation (a special nonsymmetric ARE by itself) turns out to be the unique stabilizing or almost stabilizing solution.

Keywords: quadratic matrix equation, nonsymmetric algebraic Riccati equation, M-matrices, minimal nonnegative solution, stabilizing solution, Schur method, iterative methods

1. Introduction

The main purpose of the paper is to study the quadratic matrix equation

$$X^2 - EX - F = 0, \tag{1}$$

where $E, F, X \in \mathbb{R}^{n\times n}$, E is diagonal and F is an M-matrix. Some definitions and basic results about M-matrices are given below.

For any matrices $A, B \in \mathbb{R}^{m\times n}$, we write $A \geq B (A > B)$ if $a_{ij} \geq b_{ij} (a_{ij} > b_{ij})$ for all i, j. We can then define positive matrices, nonnegative matrices, etc. A real square matrix A is called a Z-matrix if all its off-diagonal elements are

*This work was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

nonpositive. It is clear that any Z-matrix A can be written as $sI - B$ with $B \geq 0$. A Z-matrix A is called an M-matrix if $s \geq \rho(B)$, where $\rho(\cdot)$ is the spectral radius. It is called a singular M-matrix if $s = \rho(B)$; it is called a nonsingular M-matrix if $s > \rho(B)$. The spectrum of a square matrix A will be denoted by $\sigma(A)$. The open left half-plane, the open right half-plane, the closed left half-plane, and the closed right half-plane will be denoted by $\mathbb{C}_<$, $\mathbb{C}_>$, $\mathbb{C}_\leq$, and $\mathbb{C}_\geq$, respectively.

The following result is well known (see [6], for example).

Theorem 1.1: *For a Z-matrix A, the following are equivalent:*

(1) *A is a nonsingular M-matrix.*

(2) $A^{-1} \geq 0$.

(3) $Av > 0$ *for some vector* $v > 0$.

(4) $\sigma(A) \subset \mathbb{C}_>$.

The next result follows from the equivalence of (1) and (3) in Theorem 1.1.

Theorem 1.2: *Let $A \in \mathbb{R}^{n\times n}$ be a nonsingular M-matrix. If the elements of $B \in \mathbb{R}^{n\times n}$ satisfy the relations*

$$b_{ii} \geq a_{ii}, \quad a_{ij} \leq b_{ij} \leq 0, \quad i \neq j, \ 1 \leq i, j \leq n,$$

then B is also a nonsingular M-matrix.

Our study of the quadratic matrix equation (1) is motivated by noisy Wiener–Hopf problems for Markov chains.

Let Q be the Q-matrix associated with an irreducible continuous-time finite Markov chain $(X_t)_{t\geq 0}$. (A Q-matrix has nonnegative off-diagonal elements and nonpositive row sums; Q is the generator of the Markov chain.) In the study of noisy Wiener–Hopf problems for the Markov chain, we need to find, for a given diagonal matrix V and a given positive number ϵ, specific Q-matrices Γ_+ and Γ_- satisfying the quadratic matrix equations

$$\frac{1}{2}\epsilon^2 Z^2 - VZ + Q = 0 \tag{2}$$

and

$$\frac{1}{2}\epsilon^2 Z^2 + VZ + Q = 0, \tag{3}$$

respectively. In the above, V has positive *and* negative diagonal elements (This is essentially where the name Wiener–Hopf comes from.) and ϵ is the level of noise from a Brownian motion independent of the Markov chain. The solutions Γ_+ and Γ_- will be the generators of two new Markov chains. See [18, 25, 26] for more details. Rogers [25] gave a sketchy proof (using martingale theory) to show that (2) has a unique Q-matrix solution. His proof is not completely correct since, as we will see later, uniqueness fails in some cases. Nevertheless, the solutions Γ_+ and Γ_- of practical interest are still uniquely identifiable. The most efficient method so far for finding Γ_+ and Γ_- has been a diagonalization method. The validity of the method is explained by probabilistic arguments in [26] for equation (2) when the eigenvalues of the matrix

$$\begin{pmatrix} 2\epsilon^{-2}V & I \\ -2\epsilon^{-2}Q & 0 \end{pmatrix}$$

are distinct, for example.

For any Q-matrix Q, $-Q$ is a Z-matrix and $(-Q)e \geq 0$ for $e = (1, 1, \ldots, 1)^T$. So, $(\delta I - Q)e > 0$ for any $\delta > 0$. By Theorem 1.1, $\delta I - Q$ is a nonsingular M-matrix and thus $-Q$ is an M-matrix. Therefore, equations (2) and (3) are just special cases of the matrix equation (1).

We always assume that the matrices E and F are of size at least 2×2 and that $F \neq 0$. If $E = 0$ then the problem of finding a solution of (1) reduces to that of finding a square root for an M-matrix (see [1]).

We will study (1) by transforming it into an equation that belongs to a special class of nonsymmetric algebraic Riccati equations (AREs) studied earlier in [11]. Some basic results about this class of AREs are reviewed and updated in section 2. In section 3, we show that (1) has an M-matrix solution X_M when F is a nonsingular M-matrix or an irreducible singular M-matrix and that $-X_M$ is a Q-matrix if $-F$ is a Q-matrix. Several numerical methods are discussed for finding the desired M-matrix solution of (1). The methods can be used to find the solutions Γ_+ and Γ_- of the equations (2) and (3), respectively. In particular, instead of using the diagonalization method for special cases, we can use the Schur method (as in [22]) to find Γ_+ and Γ_- in the general case. Finally, section 4 contains some discussions of the general nonsymmetric AREs.

2. Results on a class of nonsymmetric AREs

For the nonsymmetric ARE

$$XCX - XD - AX + B = 0, \tag{4}$$

where A, B, C, D are real matrices of sizes $m\times m, m\times n, n\times m, n\times n$, respectively, we define two matrices H and K by

$$H = \begin{pmatrix} D & -C \\ B & -A \end{pmatrix}, \tag{5}$$

$$K = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix}. \tag{6}$$

Equation (4) in its general form has been studied in [10, 24].

In [11], we studied the ARE (4) for which K is a nonsingular M-matrix or an irreducible singular M-matrix. Nonsymmetric AREs of this type appear in transport theory (see [17]) and Wiener–Hopf factorization of Markov chains (see [25]). The solution of practical interest is the minimal nonnegative solution.

In this section, we review and update some results in [11]. All these results will be needed in the next section.

THEOREM 2.1: *If K is a nonsingular M-matrix, then* (4) *has a minimal nonnegative solution S and $D-CS$ is a nonsingular M-matrix. If K is an irreducible singular M-matrix, then* (4) *has a minimal nonnegative solution S and $D-CS$ is an M-matrix. In both cases, $Sv_1 \le v_2$, where v_1 and v_2 are positive vectors such that $K(v_1^T \;\; v_2^T)^T \ge 0$.*

Since K is a nonsingular M-matrix or an irreducible singular M-matrix, the matrices A and D are both nonsingular M-matrices. Let $\otimes$ denote the Kronecker product. Since any eigenvalue of $I \otimes A + D^T \otimes I$ is the sum of an eigenvalue of A and an eigenvalue of D (see [21], for example), it follows from the equivalence of (1) and (4) in Theorem 1.1 that $I \otimes A + D^T \otimes I$ is a nonsingular M-matrix.

Let $A = A_1 - A_2, D = D_1 - D_2$, where $A_2, D_2 \ge 0$, and A_1 and D_1 are Z-matrices. The matrix $I \otimes A_1 + D_1^T \otimes I$ is a nonsingular M-matrix by Theorem 1.2. We then have the fixed-point iteration for (4)

$$X_{k+1} = \mathcal{L}^{-1}(X_k C X_k + X_k D_2 + A_2 X_k + B), \qquad (7)$$

where the linear operator $\mathcal{L}$ is given by $\mathcal{L}(X) = A_1 X + X D_1$.

In the remainder of this section, S is always the minimal nonnegative solution of (4).

THEOREM 2.2: *For* (7) *with $X_0 = 0$, we have $X_0 \le X_1 \le \cdots$, $\lim_{k\to\infty} X_k = S$. Moreover,*

$$\begin{aligned} &\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} \\ \le\;& \rho((I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)). \end{aligned}$$

Equality holds if S is positive.

Theorems 2.4 and 2.5 below are based on corresponding results in [11] and the next more recent result (see [12]).

THEOREM 2.3: *When K is an irreducible M-matrix, we have $S > 0$ and $\hat{S} > 0$, where $\hat{S}$ is the minimal nonnegative solution of $XBX - XA - DX + C = 0$, the dual equation of* (4).

THEOREM 2.4: *If K is a nonsingular M-matrix, then the matrix H in* (5) *has n eigenvalues in $\mathbb{C}_>$ and m eigenvalues in $\mathbb{C}_<$. If K is an irreducible singular M-matrix, $u_1, u_2, v_1, v_2 > 0$ are such that $K(v_1^T \;\; v_2^T)^T = 0$ and $(u_1^T \;\; u_2^T)K = 0$, then*

(1) *If* $u_1^T v_1 = u_2^T v_2$*, then* H *has* $n-1$ *eigenvalues in* $\mathbb{C}_>$*,* $m-1$ *eigenvalues in* $\mathbb{C}_<$*, and two zero eigenvalues with only one linearly independent eigenvector. Moreover,* $D - CS$ *is an irreducible singular* M*-matrix, and* $Sv_1 = v_2$ *if* C *has no zero columns.*

(2) *If* $u_1^T v_1 > u_2^T v_2$*, then* H *has* $n-1$ *eigenvalues in* $\mathbb{C}_>$*,* m *eigenvalues in* $\mathbb{C}_<$*, and one zero eigenvalue. Moreover,* $D - CS$ *is an irreducible singular* M*-matrix, and* $Sv_1 = v_2$ *if* C *has no zero columns.*

(3) *If* $u_1^T v_1 < u_2^T v_2$*, then* H *has* n *eigenvalues in* $\mathbb{C}_>$*,* $m-1$ *eigenvalues in* $\mathbb{C}_<$*, and one zero eigenvalue. Moreover,* $D - CS$ *is an irreducible nonsingular* M*-matrix.*

THEOREM 2.5: *If* K *is a nonsingular* M*-matrix, then the matrix*

$$I \otimes (A - SC) + (D - CS)^T \otimes I \tag{8}$$

is a nonsingular M*-matrix. If* K *is an irreducible singular* M*-matrix, then* (8) *is a nonsingular (singular)* M*-matrix when* $u_1^T v_1 \neq u_2^T v_2$ $(u_1^T v_1 = u_2^T v_2)$*. Moreover,* (8) *is irreducible when* K *is an irreducible (singular or nonsingular)* M*-matrix.*

The last statement in the above theorem was not given explicitly in [11], and will be explained now. In fact, by Theorem 3.3 of [11] and Theorem 2.3 we know that $D - CS$ is an irreducible M-matrix. By taking transpose in (4), we know that S^T is the minimal nonnegative solution of the equation $YC^TY - YA^T - D^TY + B^T = 0$. Since

$$\begin{pmatrix} A^T & -C^T \\ -B^T & D^T \end{pmatrix} = \begin{pmatrix} A & -B \\ -C & D \end{pmatrix}^T$$

is also an irreducible M-matrix when K is so, it follows that $A^T - C^T S^T = (A - SC)^T$ is an irreducible M-matrix. Thus $A - SC$ is an irreducible M-matrix. Now, the irreducibility of (8) can easily be shown by a graph argument (see Theorem 2.2.7 of [6]).

3. Results on the quadratic matrix equation

In this section we will present a number of results on the equation (1), mainly by applying the results in the previous section.

3.1. The existence and uniqueness of M-matrix solutions

To apply the results of the previous section, we will change the problem of finding an M-matrix solution of (1) to that of finding a nonnegative solution of another quadratic equation (the idea was also used in [26]).

By letting $Y = \alpha I - X$, we can rewrite (1) as

$$Y^2 - Y(\alpha I) - (\alpha I - E)Y + (\alpha^2 I - \alpha E - F) = 0. \tag{9}$$

LEMMA 3.1: *If F is irreducible, then the matrix*

$$R = \begin{pmatrix} \alpha I & -I \\ -\alpha^2 I + \alpha E + F & \alpha I - E \end{pmatrix} \tag{10}$$

is also irreducible.

Proof: When F is irreducible, $-\alpha^2 I + \alpha E + F$ is also irreducible. The irreducibility of R is shown by an easy graph argument (see Theorem 2.2.7 of [6]). □

We now let the parameter α be such that

$$\alpha > 0, \quad \alpha^2 I - \alpha E - F \geq 0.$$

Let e_i and f_i be the ith diagonal element of E and F, respectively. It is easy to see that the set of all such α is the interval $[\alpha_0, \infty)$ with

$$\alpha_0 = \max_{1 \leq i \leq n} \left(e_i + \sqrt{e_i^2 + 4f_i} \right) / 2 > 0. \tag{11}$$

LEMMA 3.2: *Let $\alpha \geq \alpha_0$. If F is a nonsingular M-matrix, then R is a nonsingular M-matrix. If F is an irreducible singular M-matrix, then R is an irreducible singular M-matrix.*

Proof: Note that R is a Z-matrix. If F is a nonsingular M-matrix, then $Fv > 0$ for some $v > 0$. Take $\delta > 0$ small enough so that $Fv - \delta(\alpha^2 I - \alpha E - F)v > 0$. Direct computation shows that $R((v + \delta v)^T \ (\alpha v)^T)^T > 0$. Thus, R is a nonsingular M-matrix. If F is an irreducible singular M-matrix, then $Fv = 0$ for some $v > 0$ by the Perron-Frobenius theorem. Direct computation shows that $R(v^T \ (\alpha v)^T)^T = 0$. Thus, R is an irreducible singular M-matrix. □

We can now apply Theorems 2.1 and 2.4 to (9) and its corresponding matrices R and

$$L = \begin{pmatrix} \alpha I & -I \\ \alpha^2 I - \alpha E - F & -\alpha I + E \end{pmatrix}, \quad \alpha \geq \alpha_0. \tag{12}$$

First of all, (9) has a minimal nonnegative solution S_α when F is a nonsingular M-matrix or an irreducible singular M-matrix.

THEOREM 3.1: *If F is a nonsingular M-matrix and v is a positive vector such that $Fv \geq 0$, then L has n eigenvalues in $\mathbb{C}_>$ and n eigenvalues in $\mathbb{C}_<$, $\alpha I - S_\alpha$ is a nonsingular M-matrix, and $S_\alpha v \leq \alpha v$. If F is an irreducible singular M-matrix and u, v are positive vectors such that $Fv = 0$ and $u^T F = 0$, then*

(1) *If $u^T E v = 0$, then L has $n-1$ eigenvalues in $\mathbb{C}_>$, $n-1$ eigenvalues in $\mathbb{C}_<$, and two zero eigenvalues with only one linearly independent eigenvector, $\alpha I - S_\alpha$ is an irreducible singular M-matrix, and $S_\alpha v = \alpha v$.*

(2) *If $u^T Ev < 0$, then L has $n-1$ eigenvalues in $\mathbb{C}_>$, n eigenvalues in $\mathbb{C}_<$, and one zero eigenvalue, $\alpha I - S_\alpha$ is an irreducible singular M-matrix, and $S_\alpha v = \alpha v$.*

(3) *If $u^T Ev > 0$, then L has n eigenvalues in $\mathbb{C}_>$, $n-1$ eigenvalues in $\mathbb{C}_<$, and one zero eigenvalue, $\alpha I - S_\alpha$ is an irreducible nonsingular M-matrix, and $S_\alpha v \le \alpha v$.*

Proof: We apply Theorems 2.1 and 2.4 to (9) (so $K = R$ and $H = L$). Obviously, we can take $v_1 = v$ and $v_2 = \alpha v$ in Theorem 2.1. When F is an irreducible singular M-matrix, we can take $v_1 = v$, $v_2 = \alpha v$, $u_1^T = u^T(\alpha I - E)$, $u_2^T = u^T$ in Theorem 2.4 and thus $u_1^T v_1 - u_2^T v_2 = -u^T Ev$. The conclusions follow directly from Theorems 2.1 and 2.4. □

Since

$$\begin{pmatrix} I & 0 \\ \alpha I & I \end{pmatrix}^{-1} L \begin{pmatrix} I & 0 \\ \alpha I & I \end{pmatrix} = \begin{pmatrix} 0 & -I \\ -F & E \end{pmatrix} \equiv W, \qquad (13)$$

the matrices L and W have the same eigenvalues.

THEOREM 3.2: *If F is a nonsingular M-matrix, then (1) has exactly one M-matrix as its solution and the M-matrix is nonsingular. If F is an irreducible singular M-matrix, then (1) has M-matrix solutions and all elements of each M-matrix solution are nonzero. For case (1) or case (2) of Theorem 3.1, (1) has exactly one M-matrix as its solution and the M-matrix is singular. For case (3) of Theorem 3.1, (1) has exactly one nonsingular M-matrix as its solution but may also have singular M-matrices as its solutions.*

Proof: If F is a nonsingular M-matrix, we know from Theorem 3.1 that $\alpha I - S_\alpha$ is a nonsingular M-matrix. By the relationship between (1) and (9), $\alpha I - S_\alpha$ is a solution of (1). Note that, for any solution X of (1),

$$W \begin{pmatrix} I & 0 \\ -X & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ -X & I \end{pmatrix} \begin{pmatrix} X & -I \\ 0 & E-X \end{pmatrix}. \qquad (14)$$

Therefore, for any M-matrix solution X of (1), the column space of $(I \;\; -X^T)^T$ is the unique invariant subspace of W associated with the n eigenvalues in $\mathbb{C}_>$. The uniqueness of M-matrix solutions is thus proved.

If F is an irreducible singular M-matrix, $\alpha I - S_\alpha$ is an M-matrix and is a solution of (1). For any M-matrix solution X of (1), we can take α large enough so that $\alpha I - X$ is a nonnegative solution Y of (9). Since the minimal nonnegative solution is positive, Y is positive as well. Thus $X = \alpha I - Y$ is an M-matrix with no zero elements. For case (1) and case (2), each M-matrix solution must be singular by (14), since W has only $n-1$ eigenvalues in $\mathbb{C}_>$ by Theorem 3.1 and (13). For case (2), the uniqueness is also clear since W has only n eigenvalues in $\mathbb{C}_\ge$. For case (1), assume that (1) has two different M-matrix solutions X_1 and

X_2. We can take α large enough so that $S_\alpha^{(1)} = \alpha I - X_1$ and $S_\alpha^{(2)} = \alpha I - X_2$ are two different nonnegative solutions of (9). We may then assume that $S_\alpha^{(1)} \neq S_\alpha$. Since $S_\alpha^{(1)} \geq S_\alpha$, we have $\rho(S_\alpha^{(1)}) > \rho(S_\alpha)$ by the Perron-Frobenius theorem. However, $\alpha \geq \rho(S_\alpha^{(1)})$ since $X_1 = \alpha I - S_\alpha^{(1)}$ is an M-matrix. Thus, $\alpha > \rho(S_\alpha)$ and $\alpha I - S_\alpha$ is a *nonsingular* M-matrix solution of (1). This implies that W has at least n eigenvalues in $\mathbb{C}_>$. The contradiction proves the uniqueness for case (1). For case (3), $\alpha I - S_\alpha$ is a nonsingular M-matrix solution of (1) by Theorem 3.1. Equation (1) has no other nonsingular M-matrix solutions since W has only n eigenvalues in $\mathbb{C}_>$. However, as the next example shows, it may have singular M-matrices as its solutions. □

Example 3.1: For (1) with

$$E = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

case (3) of Theorem 3.1 happens. We find that it has two M-matrix solutions. One is nonsingular and the other is singular. The solutions (rounded to four decimals) are

$$\begin{pmatrix} 2.2703 & -1.4293 \\ -0.2703 & 0.4293 \end{pmatrix}, \quad \begin{pmatrix} 2.1715 & -2.1715 \\ -0.2890 & 0.2890 \end{pmatrix}.$$

Nevertheless, when F is an irreducible singular M-matrix and case (3) of Theorem 3.1 happens, the desired solution of (1) is the unique nonsingular M-matrix solution. This is the solution needed for later application to noisy Wiener–Hopf problems. Now, the desired M-matrix solution of (1) is uniquely identifiable for all situations in Theorem 3.1 and will be denoted by X_M. Note that $X_M = \alpha I - S_\alpha$ for all $\alpha \geq \alpha_0$. The next result follows immediately from Theorem 2.5.

PROPOSITION 3.1: *The matrix $I \otimes (X_M - E) + X_M^T \otimes I$ is a singular M-matrix for case* (1) *of Theorem* 3.1*, and is a nonsingular M-matrix for all other situations in Theorem* 3.1*. Moreover, the matrix is irreducible whenever F is irreducible.*

The following result, originally proved in [1], is a special case of Theorem 3.2 with $E = 0$.

COROLLARY 3.1: *Every nonsingular (or irreducible singular) M-matrix has a unique M-matrix square root.*

3.2. Numerical methods for finding the M-matrix solution

Recall that $X_M = \alpha I - S_\alpha$ for $\alpha \geq \alpha_0$. To find S_α, we can apply (7) to (9) with $A_2 = D_2 = 0$. By Theorem 2.2, the sequence $\{Y_k\}$ produced by the fixed-point iteration

$$Y_0 = 0, \quad Y_{k+1} = (2\alpha I - E)^{-1}(Y_k^2 + (\alpha^2 I - \alpha E - F)), \quad k \geq 0 \tag{15}$$

is monotonically increasing and converges to S_α. Moreover,

$$\limsup_{k\to\infty} \sqrt[k]{\|Y_k - S_\alpha\|} \le \rho_\alpha$$

and equality holds if S_α is positive (which is true if F is irreducible), where

$$\rho_\alpha = \rho((I\otimes(\alpha I - E) + (\alpha I)^T \otimes I)^{-1}(I\otimes S_\alpha + S_\alpha^T \otimes I)).$$

Since we have the freedom to choose $\alpha \ge \alpha_0$, we would like to make ρ_α as small as possible.

Proposition 3.2: *For case* (1) *of Theorem* 3.1, $\rho_\alpha = 1$ *for all* $\alpha \ge \alpha_0$. *For all other situations in Theorem* 3.1 *and* $\alpha_2 > \alpha_1 \ge \alpha_0$, *we have* $\rho_{\alpha_1} \le \rho_{\alpha_2} < 1$, *and* $\rho_{\alpha_1} < \rho_{\alpha_2} < 1$ *if* F *is irreducible.*

Proof: For case (1) of Theorem 3.1, $T = I\otimes(X_M - E) + X_M^T \otimes I$ is an irreducible singular M-matrix. So $Tv = 0$ for some $v > 0$. Thus,

$$(I\otimes(\alpha I - E) + (\alpha I)^T \otimes I)^{-1}(I\otimes S_\alpha + S_\alpha^T \otimes I)v = v.$$

Since $S_\alpha > 0$, it follows that $\rho_\alpha = 1$ for all $\alpha \ge \alpha_0$. For all other situations in Theorem 3.1,

$$T = (I\otimes(\alpha I - E) + (\alpha I)^T \otimes I) - (I\otimes S_\alpha + S_\alpha^T \otimes I)$$

is a regular splitting of the nonsingular M-matrix T. If F is irreducible, then T is also irreducible and thus $T^{-1} > 0$ (see Theorem 6.3.11 of [6]). The conclusions in the proposition now follow from a comparison result for regular splittings (see Theorem 3.32 or Theorem 3.36 in [27]). □

Thus, we should take $\alpha = \alpha_0$ for iteration (15). From the above result, we know that the convergence of iteration (15) is sublinear when F is an irreducible singular M-matrix and case (1) of Theorem 3.1 happens. Other solution methods are thus needed for (1). One efficient method for finding X_M is to find S_α by Newton's method:

$$(\alpha I - E - Y_i)Y_{i+1} + Y_{i+1}(\alpha I - Y_i) = \alpha^2 I - \alpha E - F - Y_i^2, \quad i = 0, 1, \ldots. \tag{16}$$

Theorem 3.3: *For the Newton iteration* (16) *with* $Y_0 = 0$, *the sequence* $\{Y_i\}$ *is well defined,* $Y_0 \le Y_1 \le \cdots$, *and* $\lim Y_i = S_\alpha$. *For case* (1) *of Theorem* 3.1, $\{Y_i\}$ *converges to* S_α *either quadratically or linearly with rate* 1/2. *For all other situations in Theorem* 3.1, $\{Y_i\}$ *converges to* S_α *quadratically.*

Proof: The proof can be completed by making some small changes in the proofs of the relevant results in section 2 of [14]. When F is a nonsingular M-matrix, we can prove by induction that, for each $k \geq 0$, $Y_k \leq S_\alpha$, $I \otimes (\alpha I - E - Y_k) + (\alpha I - Y_k)^T \otimes I$ is a nonsingular M-matrix, and $Y_k \leq Y_{k+1}$. When F is an irreducible singular M-matrix, we can prove by induction that, for each $k \geq 0$, $Y_k < S_\alpha$, $I \otimes (\alpha I - E - Y_k) + (\alpha I - Y_k)^T \otimes I$ is a nonsingular M-matrix, and $Y_k \leq Y_{k+1}$. That the matrix $I \otimes (\alpha I - E - Y_k) + (\alpha I - Y_k)^T \otimes I$ is a nonsingular M-matrix follows from the Perron-Frobenius theorem once $Y_k \leq S_\alpha$ is proved for the first case and $Y_k < S_\alpha$ is proved for the second case. □

Based on our observation on the convergence rate of the fixed-point iteration, it is advisable to take $\alpha = \alpha_0$ for Newton's method. The Sylvester equation (16) can be solved using subroutines DGEHRD, DORGHR, DHSEQR, and DTRSYL from LAPACK [2]. If the convergence of Newton's method is linear with rate 1/2 then, as explained in [13, 14], a much better approximation for S_α can be obtained by using a double Newton step, i.e., by computing $\hat{Y}_{k+1} = Y_k + 2(Y_{k+1} - Y_k) = 2Y_{k+1} - Y_k$ for a suitable k.

Another efficient method is the Schur method.

Theorem 3.4: *Let F be a nonsingular or irreducible singular M-matrix and W be the matrix in* (13). *Let U be an orthogonal matrix such that*

$$U^T W U = G$$

is a real Schur form of W, where all 1×1 or 2×2 diagonal blocks of G corresponding to eigenvalues in $\mathbb{C}_>$ appear in the first n columns of G and all 1×1 or 2×2 diagonal blocks of G corresponding to eigenvalues in $\mathbb{C}_<$ appear in the last n columns of G. If U is partitioned as

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \tag{17}$$

where $U_{11} \in \mathbb{R}^{n \times n}$, then U_{11} is nonsingular and $-U_{21}U_{11}^{-1} = X_M$, the desired M-matrix solution of (1).

Proof: Note that the column spaces of $(U_{11}^T \; U_{21}^T)^T$ and $(I \quad -X_M^T)^T$ are the same invariant subspace of W corresponding to the n eigenvalues with the largest real parts. □

If F is a nonsingular M-matrix, subroutines DGEHRD, DORGHR, DHSEQR, and DTREXC from LAPACK [2] can be used directly to compute in double precision an ordered real Schur form described in Theorem 3.4. The 1×1 or 2×2 diagonal blocks in the real Schur form obtained by DHSEQR do not generally have the required ordering. The subroutine DTREXC is used to reorder the diagonal blocks by using orthogonal transformations to interchange consecutive diagonal blocks (see [3]). To keep the number of interchanges of consecutive blocks at a minimum, we use the original ordering within the group of diagonal blocks associated with eigenvalues in $\mathbb{C}_>$ and within the group of diagonal blocks associated with eigenvalues in $\mathbb{C}_<$.

When F is an irreducible singular M-matrix and case (1) of Theorem 3.1 happens, zero is a double eigenvalue of W with index two. So, extra care should be taken to have high accuracy for the Schur method when F is an irreducible singular M-matrix. We assume that $Fv = 0$ for a known vector $v > 0$. (We have $v = e$ for noisy Wiener–Hopf problems.) We will choose the first orthogonal reduction of W by ourselves.

Let P be the Householder transformation such that $Pv = -\|v\|_2 e_1$, where $e_1 = (1, 0, \ldots, 0)^T$, and let

$$U^{(1)} = \begin{pmatrix} P & 0 \\ 0 & I \end{pmatrix}.$$

Then $(U^{(1)})^{-1} = (U^{(1)})^T = U^{(1)}$ and we have

$$U^{(1)} \begin{pmatrix} v \\ 0 \end{pmatrix} = \begin{pmatrix} -\|v\|_2 e_1 \\ 0 \end{pmatrix}.$$

Thus,

$$(U^{(1)})^T W U^{(1)} \begin{pmatrix} -\|v\|_2 e_1 \\ 0 \end{pmatrix} = (U^{(1)})^T W \begin{pmatrix} v \\ 0 \end{pmatrix} = 0.$$

So the first column of the matrix

$$N = (U^{(1)})^T W U^{(1)}$$

is zero. We can then write

$$N = \begin{pmatrix} 0 & z \\ 0 & \tilde{N} \end{pmatrix},$$

where $\tilde{N}$ is of size $(2n-1) \times (2n-1)$.

Then we let $\tilde{U}^T \tilde{N} \tilde{U} = \tilde{G}$ be an ordered real Schur form of $\tilde{N}$ obtained using the subroutines DGEHRD, DORGHR, DHSEQR, and DTREXC. We arrange the 1×1 or 2×2 diagonal blocks of $\tilde{G}$ in such a way that diagonal blocks associated with eigenvalues in $\mathbb{C}_>$ appear in the top left corner. Since one troublesome zero eigenvalue of N has been left out, we can have high accuracy for the Schur form of $\tilde{N}$. We now let

$$U^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix}.$$

By Theorem 3.1 and (13), $\tilde{G}$ has either $n-1$ or n eigenvalues in $\mathbb{C}_>$. If $\tilde{G}$ has $n-1$ eigenvalues in $\mathbb{C}_>$, then we can take $U = U^{(1)} U^{(2)}$ for Theorem 3.4. If $\tilde{G}$ has n eigenvalues in $\mathbb{C}_>$, then we use DTREXC to move the zero eigenvalue of N (or W) from the $(1,1)$ position to the $(n+1, n+1)$ position. If $u^T E v \approx 0$ then, at one stage in this process, we need to swap the zero eigenvalue with a very small positive eigenvalue. The swapping can be performed with high accuracy since we swap two 1×1 blocks (see [3]). Let $U^{(3)}$ be the orthogonal matrix used to move the zero eigenvalue to the right position. Then we can take $U = U^{(1)}U^{(2)}U^{(3)}$ for Theorem 3.4.

In Theorem 3.4, the solution X_M is found by solving $X_M U_{11} = -U_{21}$. The accuracy of X_M is thus dependent on $\kappa_2(U_{11})$, the 2-norm condition number of the matrix U_{11}. The proof of the following result is essentially the same as that of corresponding results in [16] and [20], where the results are stated for specific matrix equations.

Proposition 3.3: *Let U be an orthogonal matrix partitioned as in* (17)*. If U_{11} is nonsingular and $X = U_{21}U_{11}^{-1}$, then*

$$\kappa_2(U_{11}) \le 1 + \|X\|_2^2.$$

For the solution X_M, we have $X_M = \alpha_0 I - S_{\alpha_0}$, where α_0 is given in (11). We also have $S_{\alpha_0} v \le \alpha_0 v$ by Theorem 3.1. When the vector v is known, we can assume $v = e$ without loss of generality. More precisely, letting $V = \mathrm{diag}(v)$, we can rewrite (1) as

$$\tilde{X}^2 - \tilde{E}\tilde{X} - \tilde{F} = 0,$$

where

$$\tilde{X} = V^{-1}XV, \quad \tilde{E} = V^{-1}EV = E, \quad \tilde{F} = V^{-1}FV.$$

Notice also that α_0 will not change in the above process since $\tilde{F}$ and F have the same diagonal elements. When $v = e$, we have $S_{\alpha_0} e \le \alpha_0 e$ and thus $\|S_{\alpha_0}\|_\infty \le \alpha_0$. A very conservative estimate for $\kappa_2(U_{11})$ is thus

$$\kappa_2(U_{11}) \le 1 + \| - X_M\|_2^2 \le 1 + (\alpha_0 + \|S_{\alpha_0}\|_2)^2 \le 1 + (\alpha_0 + \sqrt{n}\,\alpha_0)^2.$$

In the study of the symmetric algebraic Riccati equation, we are interested in finding a symmetric solution X using a Schur method (see [22], for example). The symmetric solution X can be written in the form $X = U_{21}U_{11}^{-1}$. However, if U_{11} is extremely ill-conditioned the computed X can be far from symmetric. A symmetric representation for X is thus provided in [23]. In our situation, the solution X_M produced by Theorem 3.4 is supposed to be an M-matrix, but the computed X_M is not necessarily an M-matrix, particularly when the exact solution is a singular M-matrix. However, we have added an accuracy enhancement procedure in the Schur method when the method is more vulnerable to rounding errors; we have also shown that the matrix U_{11} is reasonably well-conditioned. Thus, we expect the computed X_M to have high accuracy. An easy post-processing will then (most likely) get an M-matrix approximate solution. We change any (tiny) positive off-diagonal elements of the computed X_M to zero. Moreover, when we keep a fixed number of digits for the approximate solution, we use chopping for all negative off-diagonal elements and use rounding-up for all (positive) diagonal elements. For the fixed-point iteration and Newton's method, the exact solution X_M is approximated by nonsingular M-matrices since S_α is approximated by a monotonically increasing sequence. But the above post-processing may still be used in case the M-matrix property is altered by rounding errors.

In the study of the symmetric algebraic Riccati equation, considerable effort has been made to design structure preserving methods (see [5], for example). In Theorem 3.4, the matrix W also has a special structure. This structure has been used for the first orthogonal reduction of W when F is an irreducible singular M-matrix. We do not know whether the special structure of W can be exploited any further for the Schur method. On the other hand, when $E = cI$ it is easy to see that $X_M = \frac{c}{2}I + (F + \frac{c^2}{4}I)^{1/2}$, where the square root can be found by applying a Schur method to the $n \times n$ matrix $F + \frac{c^2}{4}I$ (see [7, 15]). This yields a considerable saving of computational work. The fixed-point iteration has been simplified significantly due to the special structure of W (or the special form of the equation (1)), while the simplification in Newton's method is much less significant.

Finally, we note that the computational work for the Schur method given in Theorem 3.4 is roughly that for 4 Newton iterations or 100 fixed-point iterations.

3.3. Application to noisy Wiener–Hopf problems

If $-F$ is an irreducible Q-matrix in (1), we can take $v = e$ in Theorem 3.1. Thus, $(-X_M)e = (S_\alpha - \alpha I)e \leq 0$ and $-X_M$ is an irreducible Q-matrix. The results in section 3.1 can then be stated in terms of Q-matrices rather than M-matrices.

From Theorem 3.2 and Example 3.1, we see that *one* of the equations (2) and (3) does not necessarily have a unique Q-matrix solution. As we mentioned earlier, the matrix Q is the generator of the Markov chain $(X_t)_{t\geq 0}$. From the discussions in [18, 25] we know that the generators Γ_+ and Γ_- of two new Markov chains are obtained by finding proper Q-matrix solutions of (2) and (3). It turns out that Γ_+ (resp. Γ_-) is the unique singular Q-matrix solution when (2) (resp. (3)) has no nonsingular Q-matrix solutions. Moreover, Γ_+ (resp. Γ_-) is the unique nonsingular Q-matrix solution when (2) (resp. (3)) has singular and nonsingular Q-matrix solutions. If a Markov chain has a singular (nonsingular) Q-matrix as a generator, then the chain will live forever (die out).

We will apply the numerical methods in section 3.2 to find the matrices Γ_+ and Γ_-. We limit our attention to the more difficult case that Q is an irreducible *singular* Q-matrix (so $Qe = 0$ and $u^TQ = 0$ for some $u > 0$). This is the case of primary interest in the study of noisy Wiener–Hopf problems. It means that the original Markov chain $(X_t)_{t\geq 0}$ will live forever.

In (2) and (3), we may assume $\epsilon = \sqrt{2}$. (We can always divide the equations by the constant $\epsilon^2/2$.) Thus, with Z replaced by X, (2) and (3) are now

$$X^2 - VX + Q = 0, \tag{18}$$

and

$$X^2 + VX + Q = 0. \tag{19}$$

To find the solution Γ_+ of (18), we let $E = -V$ and $F = -Q$ in (1), and find the solution X_M of (1). Then $\Gamma_+ = -X_M$. The solution Γ_- of (19) can be found by taking $E = V$ instead.

Example 3.2: We consider (18) and (19) with

$$V = \begin{pmatrix} aI_{10} & 0 \\ 0 & bI_{10} \end{pmatrix}, \quad Q = \begin{pmatrix} -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ 1 & & & -1 \end{pmatrix} \in \mathbb{R}^{20\times 20},$$

where a and b are parameters to be specified. For this example, we have $u^T Q = 0$ for $u = e$. We consider four cases:

(a) $a = 1$, $b = -1$, so Γ_+ and Γ_- are both singular Q-matrices.

(b) $a = 2$, $b = -1$, so $\Gamma_+(\Gamma_-)$ is a singular (nonsingular) Q-matrix.

(c) $a = 2$, $b = -0.1$, so $\Gamma_+(\Gamma_-)$ is a singular (nonsingular) Q-matrix.

(d) $a = 1$, $b = -3$, so $\Gamma_+(\Gamma_-)$ is a nonsingular (singular) Q-matrix.

Table 1: Results for case (a)

	k_+	r_+	s_+	k_-	r_-	s_-
FP	567931	1.00×10^{-10}	0.25×10^{-4}	567929	1.00×10^{-10}	0.25×10^{-4}
NM	19	0.35×10^{-10}	0.14×10^{-4}	19	0.35×10^{-10}	0.14×10^{-4}
SM	–	0.35×10^{-13}	0.13×10^{-14}	–	0.74×10^{-13}	0.10×10^{-14}

Table 2: Results for case (b)

	k_+	r_+	s_+	k_-	r_-
FP	221	0.97×10^{-10}	0.42×10^{-9}	282	0.93×10^{-10}
NM	7	0.19×10^{-12}	0.83×10^{-12}	9	0.71×10^{-13}
SM	–	0.36×10^{-13}	0.18×10^{-14}	–	0.81×10^{-13}

For each case, we use the fixed-point iteration (FP), Newton's method (NM), and the Schur method (SM) to find the approximations $\tilde{\Gamma}_\pm$ for $\Gamma_\pm$. We let $r_\pm = \|(\tilde{\Gamma}_\pm)^2 \mp V\tilde{\Gamma}_\pm + Q\|_\infty$. For FP and NM, we use $\alpha = \alpha_0$, where α_0 is given in (11). For computing $\Gamma_\pm$, the two iterative methods are stopped when $r_\pm < 10^{-10}$. The number of iterations will be denoted by $k_\pm$. If $\Gamma_\pm$ is a singular Q-matrix (so

Table 3: Results for case (c)

	k_+	r_+	s_+	k_-	r_-
FP	74	0.83×10^{-10}	0.14×10^{-9}	116	0.91×10^{-10}
NM	6	0.23×10^{-13}	0.94×10^{-14}	7	0.25×10^{-13}
SM	–	0.25×10^{-13}	0.50×10^{-15}	–	0.10×10^{-12}

Table 4: Results for case (d)

	k_+	r_+	k_-	r_-	s_-
FP	260	0.97×10^{-10}	158	0.96×10^{-10}	0.27×10^{-9}
NM	7	0.31×10^{-11}	6	0.46×10^{-11}	0.17×10^{-10}
SM	–	0.11×10^{-12}	–	0.36×10^{-13}	0.30×10^{-14}

$\Gamma_\pm e = 0$), we also compute $s_\pm = \|\tilde{\Gamma}_\pm e\|_\infty$ as an additional accuracy indicator for $\tilde{\Gamma}_\pm$. Indeed, for FP and NM we have $s_\pm = \|\tilde{\Gamma}_\pm - \Gamma_\pm\|_\infty$, assuming that the monotonic convergence of FP and NM is not destroyed by rounding errors. For SM, $s_\pm$ can be arbitrarily smaller than $\|\tilde{\Gamma}_\pm - \Gamma_\pm\|_\infty$ in theory. In practice, however, it is unlikely that $s_\pm$ is smaller than $\|\tilde{\Gamma}_\pm - \Gamma_\pm\|_\infty$ by a large factor.

The numerical results are reported in Tables 1–4. For case (c), NM appears to be slightly more accurate than SM, while FP is a little less expensive than SM (but with lower accuracy). In all other cases, SM is the clear winner.

For case (a), the convergence of FP is sublinear and the convergence of NM is linear with rate 1/2. The performance of NM can be improved by using a double Newton step. If we use a double Newton step after 9 Newton iterations, we have $r_+ = 0.35 \times 10^{-10}, s_+ = 0.69 \times 10^{-10}, r_- = 0.35 \times 10^{-10}, s_- = 0.70 \times 10^{-10}$. However, case (a) causes no difficulty for SM since we separate one of the two zero eigenvalues of W in the first orthogonal reduction of W. We also note that for this case $r_\pm$ may be much smaller than $s_\pm$. Indeed, for NM (without the double Newton strategy) we typically have $r_\pm = O((s_\pm)^2)$ (see [14]). Therefore, for this case, the accuracy of the methods should be judged from $s_\pm$ rather than from $r_\pm$. The small $s_\pm$ for SM indicates that SM is very accurate even in this case.

4. Stabilizing solutions of general nonsymmetric AREs

In the previous section we have studied the quadratic matrix equation (1) by transforming it into a special nonsymmetric ARE. In this section we consider general nonsymmetric AREs. We explain why the so-called (almost) stabilizing solutions are of particular interest. It turns out that the desired solutions in sections 2 and 3 are also (almost) stabilizing solutions.

In the study of boundary value problems, we encounter the Riccati differential equation

$$\frac{dX}{dt} = XCX - XD - AX + B, \tag{20}$$

where A, B, C, D are real matrices of sizes $m \times m, m \times n, n \times m, n \times n$, respectively. See [4, 8], for example. The equilibrium solutions of (20) are the solutions of the Riccati equation (4).

Note that $\mathcal{R}(X) = XCX - XD - AX + B$ defines a mapping from $\mathbb{R}^{m\times n}$ into itself. The first Fréchet derivative of $\mathcal{R}$ at a matrix X is a linear operator $\mathcal{R}'_X : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ given by

$$\mathcal{R}'_X(Z) = -\left((A - XC)Z + Z(D - CX)\right).$$

Since the eigenvalues of the operator $\mathcal{R}'_X$ are the eigenvalues of the matrix $-(I \otimes (A - XC) + (D - CX)^T \otimes I)$, an equilibrium solution X of (20) is asymptotically stable if all eigenvalues of $I \otimes (A - XC) + (D - CX)^T \otimes I$ are in $\mathbb{C}_>$.

Definition: A solution X of (4) is called stabilizing (almost stabilizing) if all eigenvalues of $I \otimes (A - XC) + (D - CX)^T \otimes I$ are in $\mathbb{C}_>$ ($\mathbb{C}_\geq$).

Note that a stabilizing solution is also an almost stabilizing solution. When we consider (4) independent of (20), we are free to rewrite (4) as $-XCX + XD + AX - B = 0$. Thus, we could have used $\mathbb{C}_<$ ($\mathbb{C}_\leq$) in the definition of a stabilizing (almost stabilizing) solution. We remark that the stabilizing solution is also of particular interest for a nonsymmetric ARE arising in the study of spectral factorization for polynomials (see [9]).

In [19], the solutions X of (4) that are given special attention are those with all eigenvalues of $A - XC$ and $D - CX$ in $\mathbb{C}_>$ ($\mathbb{C}_<$). This implies that all eigenvalues of $I \otimes (A - XC) + (D - CX)^T \otimes I$ are in $\mathbb{C}_>$ ($\mathbb{C}_<$), but not the other way around.

It is mentioned in [22] that the Schur method can be used to find all solutions of (4). Here we present the following result about the Schur method for finding the (almost) stabilizing solution of (4).

THEOREM 4.1: *Let the eigenvalues of the matrix H in* (5) *be $\lambda_1, \lambda_2, \ldots, \lambda_{n+m}$, arranged with nonincreasing real parts. When* $\mathrm{Re}(\lambda_n) = \mathrm{Re}(\lambda_{n+1})$*, assume that* $\mathrm{Re}(\lambda_{n-1}) > \mathrm{Re}(\lambda_n)$ *and* $\mathrm{Re}(\lambda_{n+1}) > \mathrm{Re}(\lambda_{n+2})$*. Let U be an orthogonal matrix such that $U^T H U = G$ is a real Schur form of H. When* $\mathrm{Re}(\lambda_n) > \mathrm{Re}(\lambda_{n+1})$*, the diagonal blocks of G corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$ appear before the diagonal blocks of G corresponding to eigenvalues $\lambda_{n+1}, \ldots, \lambda_{n+m}$. When* $\mathrm{Re}(\lambda_n) = \mathrm{Re}(\lambda_{n+1})$*, the diagonal block(s) of G corresponding to eigenvalues λ_n and λ_{n+1} are preceded by the diagonal blocks corresponding to eigenvalues $\lambda_1, \ldots, \lambda_{n-1}$, and are followed by the diagonal blocks corresponding to eigenvalues $\lambda_{n+2}, \ldots, \lambda_{n+m}$. Then the following are true.*

(1) *Assume that* $\mathrm{Re}(\lambda_n) > \mathrm{Re}(\lambda_{n+1})$ *and U is partitioned as*

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{11} \in \mathbb{R}^{n\times n}$*. If* U_{11} *is nonsingular, then* $X = U_{21}U_{11}^{-1}$ *is the unique stabilizing solution of* (4)*. Otherwise,* (4) *has no almost stabilizing solutions.*

(2) *Assume that* $\lambda_n = \lambda_{n+1}$ *are real with only one linearly independent eigenvector and let* U *be partitioned as in part* (1)*. If* U_{11} *is nonsingular, then* $X = U_{21}U_{11}^{-1}$ *is the unique almost stabilizing solution of* (4)*. Otherwise,* (4) *has no almost stabilizing solutions.*

(3) *Assume that* λ_n *and* λ_{n+1} *are a conjugate pair corresponding to a* 2×2 *diagonal block. Then* (4) *has no almost stabilizing solutions.*

Proof: The proof is essentially the same as that of Theorem 5.2 in [11]. The positive solution there is the (almost) stabilizing solution here. Note also that we have relaxed the ordering requirement in the real Schur form of H. □

Remark: In part (2), the assumption that there is only one eigenvector associated with $\lambda_n = \lambda_{n+1}$ is essential. Without this assumption, the solution X (when it exists) is not isolated and is not stable with respect to the perturbations in the coefficient matrices in (4) (see Theorem 17.10.3 in [10]).

In view of Theorem 2.5, the minimal nonnegative solution S in section 2 is the unique stabilizing or almost stabilizing solution of (4). Also, we know from Proposition 3.1 that the solution X_M in section 3 is the unique stabilizing or almost stabilizing solution of $-X^2 + EX + F = 0$, which is equivalent to (1).

References

1. Alefeld, G. & Schneider, N. (1982). On square roots of M-matrices, *Linear Algebra Appl.* **42**: 119–132.

2. Anderson, E., Bai, Z., Bischof, C., Blackford, L. S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A. & Sorensen, D. (1999). *LAPACK Users' Guide*, 3rd ed., SIAM, Philadelphia, PA.

3. Bai, Z. & Demmel, J. W. (1993). On swapping diagonal blocks in real Schur form, *Linear Algebra Appl.* **186**: 73–95.

4. Bellman, R. & Wing, G. M. (1992). *An Introduction to Invariant Imbedding*, SIAM, Philadelphia, PA.

5. Benner, P., Mehrmann, V., & Xu, H. (1998). A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils, *Numer. Math.* **78**: 329–358.

6. Berman, A. & Plemmons, R. J. (1994). *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, PA.

7. Bjorck, A. & Hammarling, S. (1983). A Schur method for the square root of a matrix, *Linear Algebra Appl.* **52/53**: 127–140.

8. Chang, K. W. (1972). Singular perturbations of a general boundary value problem, *SIAM J. Math. Anal.* **3**: 520–526.

9. Clements, D. J. & Anderson, B. D. O. (1976). Polynomial factorization via the Riccati equation, *SIAM J. Appl. Math.* **31**: 179–205.

10. Gohberg, I., Lancaster, P. & Rodman, L. (1986). *Invariant Subspaces of Matrices with Applications*, John Wiley & Sons, Inc., New York.

11. Guo, C.-H. (2001). Nonsymmetric algebraic Riccati equations and Wiener–Hopf factorization for M-matrices, *SIAM J. Matrix Anal. Appl.* **23**: 225–242.

12. Guo, C.-H. (2001). A note on the minimal nonnegative solution of a nonsymmetric algebraic Riccati equation, Preprint. To appear in Linear Algebra Appl.

13. Guo, C.-H. & Lancaster, P. (1998). Analysis and modification of Newton's method for algebraic Riccati equations, *Math. Comp.* **67**: 1089–1105.

14. Guo, C.-H. & Laub, A. J. (2000). On the iterative solution of a class of nonsymmetric algebraic Riccati equations, *SIAM J. Matrix Anal. Appl.* **22**: 376–391.

15. Higham, N. J. (1987). Computing real square roots of a real matrix, *Linear Algebra Appl.* **88/89**: 405–430.

16. Higham, N. J. & Kim, H.-M. (2000). Numerical analysis of a quadratic matrix equation, *IMA J. Numer. Anal.* **20**: 499–519.

17. Juang, J. & Lin, W.-W. (1998). Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices, *SIAM J. Matrix Anal. Appl.* **20**: 228–243.

18. Kennedy, J. & Williams, D. (1990). Probabilistic factorization of a quadratic matrix polynomial, *Math. Proc. Cambridge Philos. Soc.* **107**: 591–600.

19. Kenney, C., Laub, A. J. & Jonckheere, E. A. (1989). Positive and negative solutions of dual Riccati equations by matrix sign function iteration, *Systems Control Lett.* **13**: 109–116.

20. Kenney, C., Laub, A. J. & Wette, M. (1989). A stability-enhancing scaling procedure for Schur–Riccati solvers, *Systems Control Lett.* **12**: 241–250.

21. Lancaster, P. & Tismenetsky, M. (1985). *The Theory of Matrices*, 2nd ed., Academic Press, Orlando, FL.

22. Laub, A. J. (1979). A Schur method for solving algebraic Riccati equations, *IEEE Trans. Automat. Control* **24**: 913–921.

23. Laub, A. J. & Gahinet, P. (1997). Numerical improvements for solving Riccati equations, *IEEE Trans. Automat. Control* **42**: 1303–1308.

24. Meyer, H.-B. (1976). The matrix equation $AZ + B - ZCZ - ZD = 0$, *SIAM J. Appl. Math.* **30**: 136–142.

25. Rogers, L. C. G. (1994). Fluid models in queueing theory and Wiener–Hopf factorization of Markov chains, *Ann. Appl. Probab.* **4**: 390–413.

26. Rogers, L. C. G. & Shi, Z. (1994). Computing the invariant law of a fluid model, *J. Appl. Probab.* **31**: 885–896.

27. Varga, R. S. (2000). *Matrix Iterative Analysis*, 2nd ed., Springer, Berlin.