SIAM J. MATRIX ANAL. APPL.
Vol. 23, No. 3, pp. 632–645

ON AN ITERATION METHOD FOR SOLVING A CLASS OF NONLINEAR MATRIX EQUATIONS*

SALAH M. EL-SAYED† AND ANDRÉ C. M. RAN‡

Abstract. This paper treats a set of equations of the form $X + A^{\star}\mathcal{F}(X)A = Q$, where $\mathcal{F}$ maps positive definite matrices either into positive definite matrices or into negative definite matrices, and satisfies some monotonicity property. Here A is arbitrary and Q is a positive definite matrix. It is shown that under some conditions an iteration method converges to a positive definite solution. An estimate for the rate of convergence is given under additional conditions, and some numerical results are given. Special cases are considered, which cover also particular cases of the discrete algebraic Riccati equation.

Key words. matrix equation, iteration methods, operator monotone functions, hermitian positive definite matrices

AMS subject classifications. 15A24, 47A62, 58F08

PII. S0895479899345571

1. Introduction. Let $\mathcal{P}(n)$ denote the set of $n \times n$ positive semidefinite matrices. We consider the following class of nonlinear matrix equations

$$X + A^{\star}\mathcal{F}(X)A = Q, \tag{1.1}$$

where $\mathcal{F}(\cdot) : \mathcal{P}(n) \to \mathcal{P}(n)$ is either monotone (meaning that $0 \leq X \leq Y$ implies that $\mathcal{F}(X) \leq \mathcal{F}(Y)$) or antimonotone (meaning that $0 \leq X \leq Y$ implies that $\mathcal{F}(X) \geq \mathcal{F}(Y)$). In particular, we shall be interested in the case where $\mathcal{F}(X)$ is generated by a function from $[0,\infty)$ to $[0,\infty)$ which is either operator monotone or operator antimonotone. For example, $\mathcal{F}(x) = x^r$ is operator monotone for $0 < r \leq 1$, while $\mathcal{F}(x) = x^{-1}$ is operator antimonotone (see, e.g., [2], where a thorough study of operator monotone functions is presented). Also, in (1.1) A is an arbitrary $n \times n$ matrix, and Q and X are in $\mathcal{P}(n)$. We also shall consider the case where $\mathcal{F}(\cdot) : \mathcal{P}(n) \to -\mathcal{P}(n)$ and is antimonotone.

In the case where $\mathcal{F}$ is monotone we shall often assume that $\mathcal{F}$, A, and Q satisfy the additional requirement that

$$A^{\star}\mathcal{F}(Q)A < Q. \tag{1.2}$$

This type of nonlinear matrix equation often arises in the analysis of ladder networks, dynamic programming, control theory, stochastic filtering, statistics and in many applications [1].

Several authors [1, 3, 4, 5, 12, 13, 17, 16] have considered such a nonlinear matrix equation problem. Compare also [15], where a different type of nonlinear matrix equation was studied. Anderson, Morley, and Trapp [1] discussed the existence of

*Received by the editors March 3, 1999; accepted for publication (in revised form) by V. Mehrmann May 30, 2001; published electronically December 14, 2001.

http://www.siam.org/journals/simax/23-3/34557.html

†Department of Mathematics, Faculty of Science, Benha University, Benha 13518, Egypt (mselsayed@frcu.eun.eg). Current address: Department of Mathematics, Scientific Departments, Education College for Girls, Al-Montazah, Buradah, Al-Qassim, Kingdom of Saudi Arabia.

‡Divisie Wiskunde en Informatica, Faculteit Exacte Wetenschappen, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands (ran@cs.vu.nl). The work of this author was partially supported by the NATO grant CRG 960700.

the positive solution to the matrix equation (1.1) when $\mathcal{F}(X) = X^{-1}$ and with the right-hand side being an arbitrary matrix, while Engwerda, Ran, and Rijkeboer [3] established and proved theorems for the necessary and sufficient conditions of existence of a positive definite solution of the matrix equation as in [1]. They discussed both the real and complex cases and established recursive algorithms to compute the largest and smallest solutions of the equation. Also Engwerda [4] proved the existence of the positive definite solution of the real matrix equation (1.1) when the right-hand side is the identity matrix, and he also found an algorithm to calculate the solution. In [12] the first author obtained necessary and sufficient conditions for existence of a positive definite solution of the matrix equation (1.1) with several forms of $\mathcal{F}(\cdot)$, without any conditions on the equation. In [8] some properties of a positive definite solution of the equation for $\mathcal{F}(X) = X^{-2}$ and with A normal were investigated. In [17, 16] several numerical algorithms for finding solutions for the case $\mathcal{F}(X) = X^{-1}$ were proposed.

The goal of this paper is to discuss the matrix equation (1.1), with general function $\mathcal{F}(X)$ which is either monotone or antimonotone. Closely connected to this equation is the map $\mathcal{G}(X) = Q - A^\star\mathcal{F}(X)A$. We shall also be interested in the dynamics of the map $\mathcal{G}$. We use iterative methods to obtain numerically a solution of the nonlinear matrix equation (1.1) under some additional conditions. In the case where $\mathcal{F} : [0,\infty) \to [0,\infty)$ is operator monotone Banach's fixed point theorem is a basic theorem to establish the existence of a positive definite solution and to obtain the rate of convergence for the sequence which is generated by an iteration method. Some numerical examples are given. For antimonotone functions a different method for proving necessary and sufficient conditions for existence of a positive definite solution is given.

The paper is organized as follows. In section 2, we discuss the monotone case. Some properties of $\mathcal{G}$ are studied. Under some conditions on $\mathcal{F}$ we obtain the rate of convergence of the iterative sequence of approximate solutions and a stopping criterion. Section 3 discusses the antimonotone case. Section 4 illustrates the performance of the method with some numerical examples and contains some remarks on the matrix equation (1.1) and on the results in the preceding sections. Section 5 discusses the equation (1.1) for maps $\mathcal{F}$ that map positive definite matrices into negative definite matrices and are antimonotone. An application to the discrete algebraic Riccati equation is given.

The following notations are used throughout the rest of the paper. The notation $A \geq 0$ ($A > 0$) means that A is positive semidefinite (positive definite), $A^\star$ denotes the complex conjugate transpose of A, and I is the identity matrix. Moreover, $A \geq B$ ($A > B$) is used as a different notation for $A - B \geq 0$ ($A - B > 0$). This induces a partial ordering on the hermitian matrices. When we say that a hermitian matrix is the smallest (largest) in some set, then this is always meant with respect to the partial ordering induced in this way. We denote by ρ the largest eigenvalue of $A^\star A$. The norm used in this paper is the spectral norm of the matrix A, i.e., $\| A \| = \sqrt{\rho(AA^\star)}$, unless otherwise noted.

Associated to (1.1) is the map $\mathcal{G}$ defined by $\mathcal{G}(X) = Q - A^\star\mathcal{F}(X)A$, which will play an important role in our analysis. Observe that a solution of (1.1) is a fixed point of $\mathcal{G}$. By $\mathcal{G}^2(X)$ we denote $\mathcal{G}(\mathcal{G}(X))$, and by $\mathcal{G}^j(X)$ the jth iterate of $\mathcal{G}$ on X.

2. The monotone case. We start with some preliminary results. We will establish and prove some theorems concerning the dynamics of $\mathcal{G}$ first. Then we shall apply Banach's fixed point theorem to obtain a positive definite solution of (1.1) under some restrictions on the map $\mathcal{F}$. Also we obtain some relations between the eigenvalues of the solution of (1.1) and the eigenvalues of the matrix A.

LEMMA 2.1. *Let $\mathcal{F}$ be monotone on $\mathcal{P}(n)$. Assume* (1.2) *holds. If X is a positive semidefinite solution of* (1.1), *then*

$$Q \geq X \geq Q - A^{\star}\mathcal{F}(Q)A = \mathcal{G}(Q). \tag{2.1}$$

In particular, X is positive definite.

Proof. From the matrix equation (1.1), we get immediately $0 \leq X \leq Q$ and $A^{\star}\mathcal{F}(X)A \leq Q$. Since X is a positive semidefinite solution of (1.1), by the monotonicity of $\mathcal{F}$ we have that $\mathcal{F}(X) \leq \mathcal{F}(Q)$. Therefore, $0 < Q - A^{\star}\mathcal{F}(Q)A \leq Q - A^{\star}\mathcal{F}(X)A = X$. □

First we show that condition (1.2) implies the existence of a fixed point of $\mathcal{G}^2$, and so implies either a periodic orbit of period 2 of the map $\mathcal{G}$ or a fixed point of $\mathcal{G}$, and gives information concerning the location of periodic orbits and, in particular, of fixed points of $\mathcal{G}$.

THEOREM 2.2. *If $\mathcal{F}$ is monotone on $\mathcal{P}(n)$ and* (1.2) *holds, then the following hold true.*

(i) *For any positive definite matrix X for which $\mathcal{G}(X)$ is positive definite we have $\mathcal{G}(Q) \leq \mathcal{G}^2(X) \leq Q$, and the set $\{X = X^{\star} \mid \mathcal{G}(Q) \leq X \leq Q\}$ is mapped into itself by $\mathcal{G}$.*

(ii) *There always exists either a periodic orbit of period 2 of the map $\mathcal{G}$ or a fixed point of $\mathcal{G}$. The sequence of matrices $\{\mathcal{G}^{(2j)}(Q)\}_{j=0}^{\infty}$ is a decreasing sequence of positive definite matrices converging to a positive definite matrix X_{∞}, and the sequence of matrices $\{\mathcal{G}^{(2j+1)}(Q)\}_{j=0}^{\infty}$ is an increasing sequence of positive definite matrices converging to a positive definite matrix $X_{-\infty}$, and the matrices $X_{\infty}, X_{-\infty}$ form either a periodic orbit of $\mathcal{G}$ of period 2, or $X_{\infty} = X_{-\infty}$, in which case it is a fixed point of $\mathcal{G}$, and hence a solution of* (1.1).

(iii) *Moreover, $\mathcal{G}$ maps the set $\{X = X^{\star} \mid X_{-\infty} \leq X \leq X_{\infty}\}$ into itself, and any periodic orbit of $\mathcal{G}$ is contained in this set. In particular, any solution of* (1.1) *is in between $X_{-\infty}$ and X_{∞}, and if $X_{-\infty} = X_{\infty}$, then there is a unique positive definite solution.*

(iv) *In the case where $X_{-\infty} = X_{\infty}$ this matrix is the global attractor for the map $\mathcal{G}$ in the following sense: for any positive definite X for which $\mathcal{G}(X)$ is positive definite as well, we have $\lim_{j\to\infty} \mathcal{G}^j(X) = X_{\infty}$.*

(v) *In the case where $X_{-\infty} \neq X_{\infty}$ the following holds: if $X \leq X_{-\infty}$, then the orbit of X under $\mathcal{G}$ converges to the periodic orbit $X_{-\infty}, X_{\infty}$ in the sense that $\lim_{j\to\infty} \mathcal{G}^{2j-1}(X) = X_{\infty}$, and $\lim_{j\to\infty} \mathcal{G}^{2j}(X) = X_{-\infty}$. If $X \geq X_{\infty}$ and $\mathcal{G}(X)$ is positive definite, then the orbit of X under $\mathcal{G}$ converges to the periodic orbit $X_{-\infty}, X_{\infty}$ in the sense that $\lim_{j\to\infty} \mathcal{G}^{2j-1}(X) = X_{-\infty}$, and $\lim_{j\to\infty} \mathcal{G}^{2j}(X) = X_{\infty}$.*

Proof. Observe that the set $[Q - A^{\star}\mathcal{F}(Q)A,\ Q] := \{X = X^{\star} | Q - A^{\star}\mathcal{F}(Q)A \leq X \leq Q\}$ is a compact and hence complete metric space. Put $\mathcal{G}(X) = Q - A^{\star}\mathcal{F}(X)A$. We observe that $\mathcal{G}$ maps the set $[Q - A^{\star}\mathcal{F}(Q)A,\ Q] = [\mathcal{G}(Q), Q]$ into itself.

Indeed, let Y be a hermitian matrix in $[Q - A^{\star}\mathcal{F}(Q)A,\ Q]$, i.e.,

$$Q - A^{\star}\mathcal{F}(Q)A \leq Y \leq Q;$$

then $\mathcal{F}(Y) \leq \mathcal{F}(Q)$ by the monotonicity, so

$$\mathcal{G}(Y) = Q - A^{\star}\mathcal{F}(Y)A \geq Q - A^{\star}\mathcal{F}(Q)A,$$

and as $\mathcal{F}(Y) \geq 0$, we have

$$\mathcal{G}(Y) = Q - A^*\mathcal{F}(Y)A \leq Q.$$

That is,

$$Q - A^*\mathcal{F}(Q)A \leq \mathcal{G}(Y) \leq Q.$$

Next we show that $\mathcal{G}$ is antimonotone on the set $[\mathcal{G}(Q), Q]$, so on this set $\mathcal{G}^2$ is monotone. Indeed, it is easily seen that $X \leq Y$ implies that

$$\mathcal{G}(X) - \mathcal{G}(Y) = A^*(\mathcal{F}(Y) - \mathcal{F}(X))A \geq 0,$$

so $\mathcal{G}$ is antimonotone. It follows that $\mathcal{G}^2$ is monotone. Now let X be any positive definite matrix. Then clearly $\mathcal{G}(X) \leq Q$. As $\mathcal{G}$ is antimonotone, this implies that $\mathcal{G}(Q) \leq \mathcal{G}^2(X) \leq Q$.

Also, for all j we have that $\mathcal{G}(Q) \leq \mathcal{G}^j(Q) \leq Q$. Taking $j = 2$ first, we see that $\mathcal{G}^2(Q) \leq Q$. Then applying $\mathcal{G}^2$ repeatedly, we see that the monotonicity of $\mathcal{G}^2$ on this set implies that the sequence $\{\mathcal{G}^{(2j)}(Q)\}_{j=0}^{\infty}$ is a decreasing sequence of positive definite matrices that is bounded below by the positive definite matrix $\mathcal{G}(Q)$. Hence it converges to a positive definite matrix X_∞, which is a fixed point of $\mathcal{G}^2$. Hence $X_\infty, \mathcal{G}(X_\infty)$ is a periodic orbit of $\mathcal{G}$ of period 2 or a fixed point.

Next take $j = 3$; then we see that $\mathcal{G}(Q) \leq \mathcal{G}^3(Q)$. Again, applying $\mathcal{G}^2$ repeatedly, the monotonicity of $\mathcal{G}^2$ on $[\mathcal{G}(Q), Q]$ implies that the sequence of matrices $\{\mathcal{G}^{(2j+1)}(Q)\}_{j=0}^{\infty}$ is an increasing sequence of positive definite matrices which is bounded above by Q. Hence this sequence has a limit $X_{-\infty}$, which is a fixed point of $\mathcal{G}^2$. Hence $X_{-\infty}, \mathcal{G}(X_{-\infty})$ is a periodic orbit of $\mathcal{G}$ of period 2 or a fixed point.

Now we shall show that $\mathcal{G}$ maps the set $[X_{-\infty}, X_\infty]$ into itself. First observe that $\mathcal{G}(Q) \leq X_\infty$, and thus, applying $\mathcal{G}^2$ repeatedly, we see that $\mathcal{G}^{(2j+1)}(Q) \leq X_\infty$ for all j. It follows that $X_{-\infty} \leq X_\infty$. Let $X_{-\infty} \leq X \leq X_\infty$. Then for all j we have $\mathcal{G}^{(2j+1)}(Q) \leq X \leq \mathcal{G}^{(2j)}(Q)$. Applying $\mathcal{G}$ and using the fact that $\mathcal{G}$ is antimonotone, we see that $\mathcal{G}^{(2j+1)}(Q) \leq \mathcal{G}(X) \leq \mathcal{G}^{(2j+2)}(Q)$. Letting $j \to \infty$ we see that $X_{-\infty} \leq \mathcal{G}(X) \leq X_\infty$.

To show that $\mathcal{G}(X_\infty) = X_{-\infty}$, observe that $X_{-\infty} \leq \mathcal{G}(X_{\pm\infty}) \leq X_\infty$. Now apply $\mathcal{G}$ to this to get $\mathcal{G}(X_\infty) \leq \mathcal{G}^2(X_{-\infty}) = X_{-\infty}$. So, $\mathcal{G}(X_\infty) = X_{-\infty}$ and $\mathcal{G}(X_{-\infty}) = X_\infty$.

Next, let $\{X_j\}_{j=1}^{p}$ be a periodic orbit of $\mathcal{G}$ of period p. Thus X_j is a fixed point of $\mathcal{G}^p$. Obviously, by part (i) we have $\mathcal{G}(Q) \leq X_j \leq Q$. Observe that $\mathcal{G}^{2p}$ is monotonic. Applying $\mathcal{G}^{2p}$ repeatedly, we readily see that $\mathcal{G}^{2kp+1}(Q) \leq X_j \leq \mathcal{G}^{2kp}(Q)$ for all $k = 0, 1, \ldots$. Hence, letting $k \to \infty$, we see that $X_{-\infty} \leq X_j \leq X_\infty$.

Finally, we shall prove (iv) and (v). Take a positive matrix X such that $\mathcal{G}(X)$ is positive definite as well. Recall that $\mathcal{G}(Q) \leq \mathcal{G}^2(X) \leq Q$. From the antimonotonicity of $\mathcal{G}$ we get that $\mathcal{G}(Q) \leq \mathcal{G}^3(X) \leq \mathcal{G}^2(Q)$. As $\mathcal{G}^2$ is monotone we deduce from these inequalities that

$$\mathcal{G}^{(2j-1)}(Q) \leq \mathcal{G}^{2j}(X) \leq \mathcal{G}^{(2j-2)}(Q),$$
$$\mathcal{G}^{(2j-1)}(Q) \leq \mathcal{G}^{(2j+1)}(X) \leq \mathcal{G}^{2j}(Q).$$

It follows that if $X_{-\infty} = X_\infty$, then $\mathcal{G}^j(X)$ converges to X_∞ as well.

In a similar way, if $X \leq X_{-\infty}$, then $\mathcal{G}(X) \geq \mathcal{G}(X_\infty) = X_\infty$. (In particular, for such X we have that $\mathcal{G}(X)$ is positive definite.) Then it follows that $\mathcal{G}(Q) \leq \mathcal{G}^2(X) \leq X_{-\infty}$. Now use the fact that $\mathcal{G}^2$ is monotone to see that $\mathcal{G}^{2j+1}(Q) \leq \mathcal{G}^{2j+2} \leq X_{-\infty}$.

As the left-hand side in these inequalities converges to $X_{-\infty}$ we see that $\mathcal{G}^{2j}(X)$ converges to $X_{-\infty}$, and hence $\mathcal{G}^{2j-1}(X)$ converges to $\mathcal{G}(X_{-\infty}) = X_{\infty}$.

Likewise, if $X \geq X_{\infty}$ and $\mathcal{G}(X)$ is positive definite, then one uses the monotonicity of $\mathcal{G}^2$ and the first part of the theorem to see that $X_{\infty} \leq \mathcal{G}^2(X) \leq Q$. Then, again using the monotonicity of $\mathcal{G}^2$ we get that $X_{\infty} \leq \mathcal{G}^{2j+2}(X) \leq \mathcal{G}^{2j}(Q)$. As the right-hand side converges to X_{∞} this proves that $\mathcal{G}^{2j}(X)$ converges to X_{∞}, and hence also $\mathcal{G}^{2j+1}(X)$ converges to $X_{-\infty}$. □

Example 2.1. As an example consider the case $\mathcal{F}(X) = X$, that is, consider the equation

$$X + A^{\star}XA = Q. \tag{2.2}$$

The condition $\mathcal{G}(Q) > 0$ now gives $Q - A^{\star}QA > 0$. From [9, Theorem 13.2.1], we see that Q being positive definite implies that A is stable with respect to the unit circle. Now consider the periodic points of $\mathcal{G}$ with period 2. They are fixed points of the equation $\mathcal{G}^2(X) = X$, which becomes

$$X = Q - A^{\star}(Q - A^{\star}XA)A = \mathcal{G}(Q) + A^{2\star}XA^2.$$

This is a standard Stein equation, and by the same theorem in [9], this has a unique solution. It follows that in the notation of the previous theorem, $X_{-\infty} = X_{\infty}$ is a fixed point of $\mathcal{G}$, and it is the unique positive definite solution to (2.2).

That the condition $\mathcal{G}(Q) > 0$ is necessary here can be seen by considering $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$. In that case $A^2 = I_2$. Taking $Q = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, we get $\mathcal{G}(Q) = 0$. This implies that *every* positive definite matrix is a solution of $\mathcal{G}^2(X) = X$; however, one easily computes that a positive definite X solves (2.2) for this case if and only if $X = \left(\begin{smallmatrix} 1 & x \\ \bar{x} & \frac{1}{2}+\mathrm{Re}\ x \end{smallmatrix}\right)$, with $|x - \frac{1}{2}| < \frac{1}{2}\sqrt{3}$.

Example 2.2. A simple example shows that the conditions of Theorem 2.2 are not sufficient to guarantee existence of a solution of (1.1). Indeed, take $n = 1$, so that we are in the scalar case, and take $Q = 4, A = 1$ and $\mathcal{F}(x) = 2$ for $x \leq 1\frac{1}{2}$ and $\mathcal{F}(x) = 3$ for $x > 1\frac{1}{2}$. Clearly there is no solution to (1.1), and a periodic orbit of $\mathcal{G}$ is given by $1, 2$.

Also, in the scalar case it is easily seen that there can be no periodic orbits of period larger than 2. To see this, one uses the fact that the real numbers are totally ordered. Indeed, let $x_1 \leq x_2 \leq \cdots \leq x_p$ be the numbers in a periodic orbit of $\mathcal{G}$, arranged in increasing order. Since $\mathcal{G}$ is antimonotone we get $\mathcal{G}(x_p) \leq \cdots \leq \mathcal{G}(x_1)$. But since this is a periodic orbit, these must be the same numbers as $x_1, \ldots, x_p$. Thus $\mathcal{G}(x_1) = x_p$ and $\mathcal{G}(x_p) = x_1$. So, x_1, x_p form a periodic orbit of period 2.

In order for a fixed point to exist, we need an additional assumption of $\mathcal{F}$, and it is natural to assume that $\mathcal{F}$ is continuous. It turns out that in this case existence of a fixed point is guaranteed.

THEOREM 2.3. *If $\mathcal{F}$ is monotone and continuous on $\mathcal{P}(n)$ and if* (1.2) *holds, then there exists a solution to* (1.1).

Proof. As we have seen, the set $[\mathcal{G}(Q), Q]$ is compact in the set of all $n \times n$ matrices. It is easily seen to be convex as well; that is, if X and Y are in between $\mathcal{G}(Q)$ and Q, then so is $tX + (1-t)Y$ for $0 \leq t \leq 1$. Indeed, $tX - t\mathcal{G}(Q)$ and $(1-t)Y - (1-t)\mathcal{G}(Q)$ are positive semidefinite, and as the set of positive semidefinite matrices is a cone, their sum is positive semidefinite as well. So $tX + (1-t)Y - \mathcal{G}(Q) \geq 0$. Likewise one shows that $Q - tX - (1-t)Y \geq 0$.

Under the condition (1.2) $\mathcal{G}$ maps this compact convex subset of the Banach space of $n \times n$ matrices into itself. Since $\mathcal{F}$ is continuous, so is $\mathcal{G}$, and hence we can apply the Schauder fixed point theorem (see, e.g., [7], section 106), to see that a fixed point of $\mathcal{G}$ must exist. $\square$

In the scalar case if $\mathcal{F}$ is continuous, then it is easily seen that there is a unique fixed point of $\mathcal{G}$, but it is not necessarily obtained as the limit of the sequence $\mathcal{G}^j(Q)$. In order for this to hold we need something additional.

In the next theorem, existence and uniqueness of a solution of (1.1) is proven, and the rate of convergence of the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^{\infty}$ is studied under an additional condition. To do this we will use Banach's fixed point theorem. Recall that a function $\mathcal{F} : [0,\infty) \to [0,\infty)$ is called *operator monotone* if for any n and any pair of hermitian $n \times n$ matrices A and B with $A \leq B$ we have $\mathcal{F}(A) \leq \mathcal{F}(B)$. See [2] for a detailed study and a complete characterization of such functions. Observe in particular that an operator monotone map is differentiable [2, Theorem V.3.6].

THEOREM 2.4. *Let $\mathcal{F} : [0,\infty) \to [0,\infty)$ be operator monotone. Let α be the smallest eigenvalue of $Q - A^\star\mathcal{F}(Q)A$ and assume that the condition* (1.2) *holds. If $q := \|A\|^2\mathcal{F}'(\alpha) < 1$, then* (1.1) *has a unique positive solution X_∞ and the iteration $X_{n+1} = Q - A^\star\ \mathcal{F}(X_n)\ A$, started at $X_0 = Q$, converges to X_∞ with*

$$\|X_{j+1} - X_j\| \ \leq\ q\ \|X_j - X_{j-1}\|. \tag{2.3}$$

Moreover, there are no periodic orbits of $\mathcal{G}$, and the iteration process converges to X_∞ from any positive definite X_0 for which $Q - A^\star\mathcal{F}(X_0)A$ is positive definite.

Proof. Observe that the set $[Q - A^\star\mathcal{F}(Q)A,\ Q] = [\mathcal{G}(Q), Q]$ is a compact and hence complete metric space and that $\mathcal{G}$ maps this set into itself. We shall prove that the operator $\mathcal{G}$ is a strict contraction on the set $[Q - A^\star\mathcal{F}(Q)A,\ Q]$. For this purpose, let X and Y be in $[Q - A^\star\mathcal{F}(Q)A,\ Q]$. Then

$$\|\mathcal{G}(X) - \mathcal{G}(Y)\| = \|A^\star\ (\mathcal{F}(X) - \mathcal{F}(Y))\, A\| \leq \|A\|^2\|\mathcal{F}(X) - \mathcal{F}(Y)\|. \tag{2.4}$$

Since X and Y both are greater than or equal to $Q - A^\star\mathcal{F}(Q)A > 0$, we have

$$X \geq \alpha I \qquad \text{and} \qquad Y \geq \alpha I\ (\alpha > 0).$$

(Recall that α is the smallest eigenvalue of $Q - A^\star\mathcal{F}(Q)A$, which is positive by assumption.) So we have by Theorem X.3.8 in [2] that

$$\|\mathcal{F}(X) - \mathcal{F}(Y)\| \leq \mathcal{F}'(\alpha)\|X - Y\|. \tag{2.5}$$

By combining the two inequalities (2.4) and (2.5), we get

$$\|\mathcal{G}(X) - \mathcal{G}(Y)\| \leq \|A\|^2\ \mathcal{F}'(\alpha)\|X - Y\| = q\ \|X - Y\|. \tag{2.6}$$

Since $q < 1$ by assumption, we can apply Banach's fixed point theorem; hence (1.1) has a unique positive solution X_∞ in $[Q - A^\star\mathcal{F}(Q)A,\ Q]$. By Lemma 5.5 it follows that X_∞ is the unique positive definite solution. Moreover, the sequence of successive approximations

$$X_{j+1} = Q - A^\star\ \mathcal{F}(X_j)\ A = \mathcal{G}(X_j), \quad j = 0, 1, 2, \ldots,$$

started at $X_0 = Q$, i.e., $X_j = \mathcal{G}^j(Q)$, converges to X_∞.

As $X_{-\infty} = X_\infty$ there can be no periodic orbits of $\mathcal{G}$, and the convergence of $\mathcal{G}^j(X_0)$ to X_∞ for any X_0 for which $\mathcal{G}(X_0)$ is positive definite follows from Theorem 2.2. This completes the proof of the theorem. $\square$

COROLLARY 2.5. *If all assumptions in the above theorem are satisfied and $X_0 = Q$, then*

$$\|X_{j+1} - X_j\| \le q^j \|X_1 - X_0\| = q^j \|A^\star \mathcal{F}(Q) A\|. \tag{2.7}$$

It follows from this that if $q < 1$ we have the following error bound:

$$\|X_\infty - X_j\| \le q^j \|A^\star \mathcal{F}(Q) A\|.$$

Indeed, recall that X_∞ is always between two consecutive elements of the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$.

The next corollary describes the number of iterations to be taken to ensure that $\|X_\infty - X_j\| \le \varepsilon$.

COROLLARY 2.6. *If ε is a convergence tolerance and $X_0 = Q$, then the number n of iterations to be taken is at most*

$$n = \left\lceil \frac{\ln \varepsilon - \ln \|A^\star \mathcal{F}(Q)\ A\|}{\ln q} \right\rceil + 1.$$

In the theory of Stein equations, i.e., equations of the form $X - A^\star X A = Q$, with $Q > 0$, there are well-known results relating the eigenvalues of A and X. The following theorem may be viewed as an analogue of these results for the type of equations under consideration in this section.

THEOREM 2.7. *Let $\mathcal{F} : [0,\infty) \to [0,\infty)$ be operator monotone. Let X be a positive definite solution of* (1.1), *and denote by μ_+ and μ_- the largest and smallest eigenvalue of X, respectively. Also, denote by q_+ and q_- the largest and smallest eigenvalue of $Q - X$, respectively. If λ is an eigenvalue of A, then*

$$\sqrt{\frac{q_-}{\mathcal{F}(\mu_+)}} \le |\lambda| \le \sqrt{\frac{q_+}{\mathcal{F}(\mu_-)}}.$$

In the particular case where $Q = I$, then

$$\sqrt{\frac{1-\mu_+}{\mathcal{F}(\mu_+)}} \le |\lambda| \le \sqrt{\frac{1-\mu_-}{\mathcal{F}(\mu_-)}}.$$

Proof. Let v be an eigenvector corresponding to an eigenvalue λ of the matrix A and $\|v\| = 1$. Then, with the usual scalar product denoted by $\langle \cdot, \cdot \rangle$, we have

$$\langle Xv, v\rangle + \langle A^\star \mathcal{F}(X) A v, v\rangle = \langle Qv, v\rangle.$$

So

$$|\lambda|^2 \langle \mathcal{F}(X) v, v\rangle = \langle (Q - X) v, v\rangle.$$

Now $Q - X$ is positive definite and $q_- I \le Q - X \le q_+ I$, and the largest eigenvalue of $\mathcal{F}(X)$ is $\mathcal{F}(\mu_+)$, so

$$|\lambda|^2 \mathcal{F}(\mu_+) \ge q_-.$$

Likewise, as the smallest eigenvalue of $\mathcal{F}(X)$ is $\mathcal{F}(\mu_-)$, we have that

$$|\lambda|^2 \mathcal{F}(\mu_-) \le q_+.$$

In the case where $Q = I$, simply observe that $q_- = 1 - \mu_+$ and $q_+ = 1 - \mu_-$. ∎

3. The antimonotone case. In this section $\mathcal{F}$ is assumed to be antimonotone. First, we show that this implies that $\mathcal{G}$, defined by $\mathcal{G}(X) = Q - A^\star\mathcal{F}(X)A$, is monotone. Indeed, let $0 < X \leq Y$. Then

$$\mathcal{G}(Y) - \mathcal{G}(X) = A^\star(\mathcal{F}(X) - \mathcal{F}(Y))A,$$

and as $\mathcal{F}$ is antimonotone, the latter is positive semidefinite. So $\mathcal{G}(X) \leq \mathcal{G}(Y)$. Next we present necessary and sufficient conditions for the existence of a positive definite solution.

THEOREM 3.1. *Let $\mathcal{F}$ be antimonotone on $\mathcal{P}(n)$. There is a positive definite solution to* (1.1) *if and only if the sequence of matrices $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is positive definite for all j, and the sequence $\{(\mathcal{G}^j(Q))^{-1}\}_{j=0}^\infty$ is uniformly bounded.*

In this case the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is decreasing and converges to the largest positive definite solution of (1.1).

Proof. Suppose that X_0 is an arbitrary positive definite solution of (1.1). Clearly $X_0 \leq Q$. As $\mathcal{G}$ is monotone it follows that $\mathcal{G}^j(Q) \geq X_0$ for all j. Thus $\mathcal{G}^j(Q)$ is positive definite for all j. As $Q - A^\star\mathcal{F}(Q)A = \mathcal{G}(Q) \leq Q$ and $\mathcal{G}$ is monotone we see that the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is a decreasing sequence. As this sequence is bounded below by X_0 it converges to a positive definite solution of (1.1), which we denote by X_∞. Observe that it may be the case that $X_0 \neq X_\infty$, but certainly $X_\infty \geq X_0$. This also proves that X_∞ is the largest positive definite solution. Then $(\mathcal{G}^j(Q))^{-1}$ converges to X_∞^{-1} and hence is uniformly bounded.

Conversely, assume that $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is positive definite for all j, and the sequence $\{(\mathcal{G}^j(Q))^{-1}\}_{j=0}^\infty$ is uniformly bounded. We have already seen that the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is decreasing. As each element in the sequence is positive definite, it is also bounded below (by the zero matrix). Thus there exists a limit, again denoted by X_∞, which is positive semidefinite. We only have to show that X_∞ is invertible, as it then will follow that X_∞ solves (1.1). Since $\{\mathcal{G}^j(Q)\}_{j=0}^\infty$ is a decreasing sequence of positive definite matrices it follows that $\{(\mathcal{G}^j(Q))^{-1}\}_{j=0}^\infty$ is an increasing sequence of positive definite matrices. As this sequence is uniformly bounded, it has a limit, say Y_∞, which is positive definite. Then clearly $Y_\infty^{-1} = X_\infty$, which therefore is also positive definite. $\square$

Concerning the order of convergence we can be less explicit here than in the previous section. In fact, in order to determine the order of convergence we would like to have an a priori lower bound on the eigenvalues of X_∞, as can be seen from the following theorem.

THEOREM 3.2. *Let $\mathcal{F} : [0, \infty) \to [0, \infty)$ be operator antimonotone. Assume that* (1.1) *has a positive definite solution. Let β be less than or equal to the smallest eigenvalue of the largest positive definite solution X_∞. Put $q = |\mathcal{F}'(\beta)| \cdot \|A\|^2$. Also denote $X_j = \mathcal{G}^j(Q)$ for $j = 0, 1, \ldots$. Then for $j \geq 1$*

$$\|X_{j+1} - X_j\| \leq q\|X_j - X_{j-1}\|.$$

Proof. The proof uses the same methods as the proof of Theorem 2.2. In fact, we know from the proof of the previous theorem that $X_\infty \leq X_j \leq Q$ for all j. Now if $X_\infty \leq Y \leq Q$, then $X_\infty \leq \mathcal{G}(Y) \leq X_1 \leq Q$, and for X and Y in $[X_\infty, Q]$ we have

$$\|\mathcal{G}(X) - \mathcal{G}(Y)\| \leq \|A\|^2\|\mathcal{F}(X) - \mathcal{F}(Y)\|.$$

As X and Y both are greater than or equal to X_∞ we have that $X \geq \beta I$ and $Y \geq \beta I$. Then we again apply Theorem X.3.8 in [2] to finish the proof. $\square$

In the case where A is invertible and $\mathcal{F}^{-1}$ exists and is antimonotone as well, we can find an a priori lower bound for X_∞ as follows. Suppose X is any positive definite solution; then $A^\star \mathcal{F}(X) A \leq Q$, from which it follows that $X \geq \mathcal{F}^{-1}(A^{-\star} Q A^{-1})$. Thus in that case we can take β to be the smallest eigenvalue of $\mathcal{F}^{-1}(A^{-\star} Q A^{-1})$. For instance, in the case where $\mathcal{F}(x) = x^{-1}$ and A is invertible, we can use such an estimate.

4. Remarks and numerical results. So far we considered general nonlinear matrix equations and achieved general conditions for the existence of a positive definite solution. Moreover, we discussed an iterative algorithm from which a solution can always be calculated numerically whenever the equation is solvable.

Let us see how this works in particular cases. As a first example, take $\mathcal{F}(x) = x^r$, with $0 < r < 1$, take $Q = I$, and take A a contraction. Then all conditions of Theorem 2.2 are satisfied. The condition $q < 1$ of Theorem 2.4 becomes $\rho r(1-\rho)^{r-1} < 1$, where ρ denotes $\|A\|^2$.

As a second example, take $\mathcal{F}(x) = \frac{1}{x}$. Then we can apply the results of section 3. Assuming that A is invertible, we can take for β the minimal eigenvalue of $AQ^{-1}A^\star$. In the case in which we take $Q = I$, this is the minimal eigenvalue of $AA^\star$. We get that $\mathcal{F}'(\beta) = -\frac{1}{\beta^2}$, so $q = \frac{\rho}{\beta^2}$. If we would like $q < 1$, then that amounts to $\rho < \beta^2$. Observe that for this choice of β, however, we always have that $\rho > \beta$.

Finally, we note that the results obtained so far on the iterative procedure for finding positive definite solutions are more general than those obtained in [3, 4, 12, 16, 17] in the sense that we deal with a larger class of matrix equations. It should be emphasized that the methods proposed in [6, 16, 17], while performing probably better for the special case under consideration there ($\mathcal{F}(X) = X^{-1}$), may not be so readily applied to the very general case we have under consideration here. The reader is referred to [6, 16, 17], where the numerical procedures are discussed and calculated in greater detail.

In the remainder of this section, we report some numerical results. These numerical results describe the performance of the algorithm. The numerical experiments were carried out on an IBM-PC Pentium 233 MHz computer with double precision. The machine precision is approximately 1.11×10^{-16}. We use the FORTRAN language with FORTRAN PowerStation (visual workbench version 1.00) to calculate the appended results. Table 1 indicates the convergence pattern of the iterative sequence of approximate solutions. In the example we take $Q = I$. In Table 1 n denotes the order of the matrix, k denotes the number of iterations, ε_k denotes $\|X_k + A^\star \mathcal{F}(X_k) A - I\|_\infty$, and R_k denotes the relative errors $R_k = \frac{\|X_\infty - X_k\|_\infty}{\|X_\infty\|_\infty}$, where X_∞ is taken to be the final iterate after $\varepsilon_k < 10^{-8}$ is satisfied.

The algorithm has been tested for one form of $\mathcal{F}(X)$. We take $\mathcal{F} = \sqrt{X}$ (the operator monotone case). Observe that here we only comment on the iterative procedure described in the present paper. Although this works fine for many cases, there is no claim that it is the best available procedure.

Example 4.1.

$$A = \frac{0.5\, B}{\|B\|_\infty}, \quad \text{where } B = (b_{ij}) = (i + j + 1). \tag{4.1}$$

In this example $\rho = 0.25$ and $q = 0.144338 < 1$.

In Table 1, the number of iterations can be expected if we use Corollary 2.6. In the example the number of iterations is at least 8 ($\|X_\infty - X_j\| < 10^{-8}$). The results show that with this method the efficiency and the accuracy achieved are acceptable

TABLE 1
Error analysis for (4.1).

n	k	ε_k	R_k
4	13	1.122320E-08	1.745058E-09
6	11	7.450581E-09	1.042610E-08
8	10	1.303852E-08	1.253618E-08
10	10	3.725290E-09	4.896479E-09
12	10	3.725290E-09	4.818559E-09
14	9	6.519258E-09	8.335872E-09
16	9	2.793968E-09	3.541677E-09
18	9	1.862645E-09	2.345221E-09
20	9	1.396984E-09	1.166283E-09
22	8	9.778887E-09	1.045015E-08
24	8	8.676356E-09	6.984919E-09

in the sense that we get a numerically reliable solution (in single precision) within a relatively small number of iterations.

5. Antimonotone $\mathcal{F}$ mapping positive definite to negative definite matrices. In this section we consider (1.1) under the assumption that $\mathcal{F} : \mathcal{P}(n) \to -\mathcal{P}(n)$ is antimonotone. Obviously, this implies that $\mathcal{G}$ is a monotone map mapping positive definite matrices into matrices that are greater than or equal to Q.

First we state a general theorem concerning this class of equations.

THEOREM 5.1. *Let $\mathcal{F} : \mathcal{P}(n) \to -\mathcal{P}(n)$ be antimonotone. Assume there is a positive definite matrix $\tilde{X}_0$ such that $\tilde{X}_0 \geq \mathcal{G}(\tilde{X}_0)$. Then the following hold true.*

(i) *The sequence $\mathcal{G}^j(Q)$ is an increasing sequence that is bounded above. Its limit, which we denote by X_-, is a solution to* (1.1).

(ii) *The sequence $\mathcal{G}^j(\tilde{X}_0)$ is a decreasing sequence that is bounded below. Its limit is a solution to* (1.1).

(iii) *X_- is the smallest positive definite solution to* (1.1). *Moreover, if $X_j, j = 1, \ldots, p$ is a periodic orbit of $\mathcal{G}$ consisting of positive definite matrices, then $X_j \geq X_-$ for all j.*

Proof. Observe that for any positive definite matrix X we have $\mathcal{G}(X) \geq Q$ as $\mathcal{F}$ maps positive definite matrices into negative definite matrices. Now $\tilde{X}_0 \geq \mathcal{G}(\tilde{X}_0) \geq Q$. Since $\mathcal{G}$ is monotonic, repeated application of $\mathcal{G}$ gives

$$\tilde{X}_0 \geq \mathcal{G}^j(\tilde{X}_0) \geq \mathcal{G}^{j+1}(\tilde{X}_0) \geq \mathcal{G}^j(Q) \geq Q.$$

This proves (i) and (ii).

To prove (iii), suppose $X_j, j = 1, \ldots, p$, is a periodic orbit of $\mathcal{G}$ of period p. Then X_j is a fixed point of $\mathcal{G}^p$. In particular $X_j \geq Q$, and then by monotonicity of $\mathcal{G}^p$ we get that $X_j \geq \mathcal{G}^{kp}(Q)$ for all positive integers k. Thus $X_- \leq X_j$. □

Part of the theorem can be restated as follows.

THEOREM 5.2. *Let $\mathcal{F} : \mathcal{P}(n) \to -\mathcal{P}(n)$ be antimonotone. Assume that there is a solution $\tilde{X}_0$ to the inequality*

$$X + A^{\star}\mathcal{F}(X)A \geq Q.$$

Then there is a solution to the equation

$$X + A^{\star}\mathcal{F}(X)A = Q,$$

and the sequence $\{\mathcal{G}^j(Q)\}_{j=0}^{\infty}$ increases to the smallest hermitian solution of this equation.

For an important example of this type of map consider the discrete algebraic Riccati equation

$$X = A^\star XA + Q - (A^\star XB + S)(R + B^\star XB)^{-1}(B^\star XA + S^\star),$$

where we assume that $Q = Q^\star$ and $R = R^\star$ is invertible. In linear quadratic optimal control problems usually $R > 0$ and $Q - SR^{-1}S^\star \geq 0$. It is also well known that if in addition $(A, Q - SR^{-1}S^\star)$ is observable, the solution of interest is positive definite; see [10]. By Proposition 12.1.1 in [10] we can restrict our attention to the case $S = 0$, i.e., to the equation

$$X = A^\star XA + Q - A^\star XB(R + B^\star XB)^{-1}B^\star XA, \tag{5.1}$$

where we assume that $Q \geq 0$ and R is positive definite. A good source of information concerning the discrete algebraic Riccati equation is [10], where iteration methods for solving it are also discussed. See also [14].

Here we shall restrict our attention to the particular case where Q is positive definite. It should be noted that this is a serious restriction, as most practical applications in linear quadratic optimal control theory would have a Q which is not invertible. However, here we are interested to see how the methods developed before can be applied to (5.1). In the meantime we develop results for several wider classes of equations, as will be seen in what follows.

For this equation we first reduce to the case where $R = I$ by replacing B by $BR^{-\frac{1}{2}}$. Then it is of the form (1.1) with

$$\mathcal{F}(X) = -X + XB(I + B^\star XB)^{-1}B^\star X.$$

We shall show that $\mathcal{F}$ maps positive definite matrices to negative definite matrices and is antimonotone. Indeed, first observe that

$$B(I + B^\star XB)^{-1} = (I + BB^\star X)^{-1}B,$$

so that

$$\begin{aligned}\mathcal{F}(X) &= -X + X(I + BB^\star X)^{-1}BB^\star X\\ &= -X + X(I + BB^\star X)^{-1}\{(I + BB^\star X) - I\}\\ &= -X(I + BB^\star X)^{-1}.\end{aligned}$$

For X invertible we get

$$\mathcal{F}(X) = -(X^{-1} + BB^\star)^{-1}.$$

Clearly it follows that $\mathcal{F}$ maps $\mathcal{P}(n)$ into $-\mathcal{P}(n)$. Furthermore, if $X \geq Y > 0$ then $0 < X^{-1} + BB^\star \leq Y^{-1} + BB^\star$, and hence $(X^{-1} + BB^\star)^{-1} \geq (Y^{-1} + BB^\star)^{-1}$, so that $\mathcal{F}$ is antimonotone.

Compare Theorem 5.2 to [11, Theorem 3.1]. There, a similar statement was proved concerning the discrete algebraic Riccati equation. It should be noted, however, that the result in [11] does not follow from the above theorem.

Let us see how we can apply Theorem 5.1 to the case of the discrete algebraic Riccati equation (5.1). We shall only consider a special case, namely the case where, as before, $R = I$ and Q is positive definite, but in addition we require A to be stable. First we observe that in that case $\mathcal{F}(X) = -X + \mathcal{H}(X)$, where $\mathcal{H} : \mathcal{P}(n) \to \mathcal{P}(n)$. That situation is treated in the following theorem.

THEOREM 5.3. *Let $\mathcal{F} : \mathcal{P}(n) \to -\mathcal{P}(n)$ be antimonotone. Assume $\mathcal{F}(X) = -X + \mathcal{H}(X)$, where $\mathcal{H} : \mathcal{P}(n) \to \mathcal{P}(n)$, and that A is stable. Denote by $\tilde{X}_0$ the unique solution to the Stein equation*

$$X - A^\star XA = Q. \tag{5.2}$$

Then the sequence $\{\mathcal{G}^j(\tilde{X}_0)\}$ is a decreasing sequence of positive definite matrices having a positive definite limit X_+, and X_+ is the largest positive definite solution of

$$X + A^\star \mathcal{F}(X)A = Q.$$

Proof. Observe that $\mathcal{G}(\tilde{X}_0) = Q - A^\star\mathcal{F}(\tilde{X}_0)A = Q + A^\star\tilde{X}_0A - A^\star\mathcal{H}(\tilde{X}_0)A = \tilde{X}_0 - A^\star\mathcal{H}(\tilde{X}_0)A \leq \tilde{X}_0$. So we can apply Theorem 5.1 to see that the sequence $\{\mathcal{G}^j(\tilde{X}_0)\}$ is a decreasing sequence having a positive definite limit X_+, which is a solution to (1.1). So we only have to show that it is the largest positive definite solution.

Let X be any positive definite solution to (1.1). Then

$$\begin{aligned}\tilde{X}_0 - X &= Q + A^\star\tilde{X}_0A - (Q - A^\star\mathcal{F}(X)A)\\ &= A^\star(\tilde{X}_0 + \mathcal{F}(X))A\\ &= A^\star(\tilde{X}_0 - X)A + A^\star\mathcal{H}(X)A.\end{aligned}$$

So $\tilde{X}_0 - X$ solves the Stein equation

$$\tilde{X}_0 - X - A^\star(\tilde{X}_0 - X)A = A^\star\mathcal{H}(X)A.$$

As $\mathcal{H}(X)$ is positive definite and A is stable we see that $\tilde{X}_0 - X$ is positive semidefinite. So $X \leq \tilde{X}_0$. As $\mathcal{G}$ is monotone, this implies that $X = \mathcal{G}^j(X) \leq \mathcal{G}^j(\tilde{X}_0)$ for all positive integers j, so that $X \leq X_+$. □

Observe that this result can be applied directly to the discrete algebraic Riccati equation under the assumptions that Q is positive definite and A is stable. (Recall that we may assume that $R = I$ without loss of generality.) That yields the following result.

COROLLARY 5.4. *Let A be stable and let Q be positive definite. Let $\tilde{X}_0$ be the unique solution of the Stein equation* (5.2). *Define the sequence of matrices $\{X_j\}$ by*

$$X_{j+1} = Q + A^\star X_jA - A^\star X_jB(R + B^\star X_jB)^{-1}B^\star X_jA,$$

with $X_0 = \tilde{X}_0$. Then this sequence of matrices decreases to the largest positive definite solution of (5.1).

Define the sequence of matrices $\{Q_j\}$ by

$$Q_{j+1} = Q + A^\star Q_jA - A^\star Q_jB(R + B^\star Q_jB)^{-1}B^\star Q_jA,$$

with $Q_0 = Q$. Then this sequence of matrices increases to the smallest positive definite solution of (5.1).

This result is well known—as a matter of fact it can be proven that under the present conditions there is a unique positive definite solution (see, e.g., [10, Theorem 13.5.3]). To obtain this result we need to use much more of the structure of the map $\mathcal{F}$. We start with a lemma.

LEMMA 5.5. *Assume that $\mathcal{F} : \mathcal{P}(n) \to -\mathcal{P}(n)$ is antimonotone and is of the form $\mathcal{F}(X) = -X + X\mathcal{H}(X)X$, where $\mathcal{H}$ satisfies $\mathcal{H}(X)X\mathcal{H}(X) \leq \mathcal{H}(X)$. Then for* *every positive definite solution X of $X = \mathcal{G}(X)$ the matrix A_X defined by $A_X = A - \mathcal{H}(X)XA$ is stable.*

Proof. Let X be a positive definite solution of (1.1) and compute

$$\begin{aligned} & X - A_X^\star X A_X \\ & = X - A^\star XA + 2A^\star X\mathcal{H}(X)XA - A^\star X\mathcal{H}(X)X\mathcal{H}(X)XA \\ & = Q + A^\star X\mathcal{H}(X)XA - A^\star X\mathcal{H}(X)X\mathcal{H}(X)XA. \end{aligned}$$

From the assumption on $\mathcal{H}$ it follows that $X - A_X^\star XA_X$ is positive definite. As X is positive definite we get that A_X is stable (see, e.g., [9, section 13.2]). □

The next theorem describes conditions which are satisfied in the case of the discrete algebraic Riccati equation and which allow us to deduce that there is a unique positive definite solution.

THEOREM 5.6. *Let $\mathcal{F}(X) = -X + X\mathcal{H}(X)X$ map positive definite matrices into negative definite matrices, and let it be antimonotone. Assume that $\mathcal{H}$ satisfies the following two properties:*

$$\mathcal{H}(X)X\mathcal{H}(X) \leq \mathcal{H}(X), \tag{5.3}$$

$$\mathcal{H}(Y) - \mathcal{H}(X) = \mathcal{H}(X)(X - Y)\mathcal{H}(Y). \tag{5.4}$$

Then there is a unique positive definite solution to the equation

$$X - A^\star XA + A^\star X\mathcal{H}(X)XA = Q. \tag{5.5}$$

Proof. From the results obtained so far it follows that we only have to show that $X_+ = X_-$. To do this, put $A_\pm = A - \mathcal{H}(X_\pm)X_\pm A$. As (5.3) holds, we can apply the previous lemma to see that both these matrices are stable. Now compute

$$\begin{aligned} & X_+ - X_- - A_+^\star(X_+ - X_-)A_- \\ & = X_+ - X_- - (A^\star - A^\star X_+\mathcal{H}(X_+))X_+(A - \mathcal{H}(X_-)X_-A) \\ & \quad + (A^\star - A^\star X_+\mathcal{H}(X_+))X_-(A - \mathcal{H}(X_-)X_-A) \\ & = X_+ - A^\star X_+A + A^\star X_+\mathcal{H}(X_+)X_+A \\ & \quad + A^\star X_+\mathcal{H}(X_-)X_-A - A^\star X_+\mathcal{H}(X_+)X_+\mathcal{H}(X_-)X_-A \\ & \quad - X_- + A^\star X_-A - A^\star X_-\mathcal{H}(X_-)X_-A \\ & \quad - A^\star X_+\mathcal{H}(X_+)X_-A + A^\star X_+\mathcal{H}(X_+)X_-\mathcal{H}(X_-)X_-A. \end{aligned}$$

Using the fact that X_+ and X_- are both solutions to (5.5) we see that

$$\begin{aligned} & X_+ - X_- - A_+^\star(X_+ - X_-)A_- \\ & = A^\star X_+(\mathcal{H}(X_-) - \mathcal{H}(X_+) + \mathcal{H}(X_+)X_-\mathcal{H}(X_-) - \mathcal{H}(X_+)X_+\mathcal{H}(X_-))X_-A = 0, \end{aligned}$$

where the last equality follows from (5.4). As A_+ and A_- are both stable, the equation $Y - A_+^\star YA_- = 0$ has a unique solution, being the zero matrix. It follows that $X_+ = X_-$, as desired. □

It is easily seen in the case of the discrete algebraic Riccati equation that both conditions (5.3) and (5.4) are satisfied. Indeed, in that case $\mathcal{H}(X) = B^\star(R + B^\star XB)^{-1}B$, where we may assume without loss of generality that $R = I$, as before. Then

$$\begin{aligned}&\mathcal{H}(X)X\mathcal{H}(X)\\&= B(I+B^\star XB)^{-1}B^\star XB(I+B^\star XB)^{-1}B^\star\\&= B(I+B^\star XB)^{-1}B^\star - B(I+B^\star XB)^{-2}B^\star\\&\le B(I+B^\star XB)^{-1}B^\star = \mathcal{H}(X).\end{aligned}$$

Also,

$$\begin{aligned}&\mathcal{H}(Y)-\mathcal{H}(X)\\&= B\{(I+B^\star YB)^{-1}-(I+B^\star XB)^{-1}\}B^\star\\&= B(I+B^\star XB)^{-1}\{(I+B^\star XB)-(I+B^\star YB)\}(I+B^\star YB)^{-1}B^\star\\&= B(I+B^\star XB)^{-1}B^\star(X-Y)B(I+B^\star YB)^{-1}B^\star\\&= \mathcal{H}(X)(X-Y)\mathcal{H}(Y).\end{aligned}$$

Thus the theorem above can be applied directly to the discrete algebraic Riccati equation.

COROLLARY 5.7. *Assume that Q is positive definite and that A is a stable matrix. Then the algebraic Riccati equation* (5.1) *has a unique positive definite solution, which is the stabilizing solution.*

Acknowledgment. The authors would like to thank Prof. Salah E. El-Gendi for his careful reading of one of the early versions of the manuscript.

REFERENCES

[1] W. N. ANDERSON, JR., T. D. MORLEY, AND G. E. TRAPP, *Positive solution to $X = A - BX^{-1}B^\star$*, Linear Algebra Appl., 134 (1990), pp. 53–62.

[2] R. BHATIA, *Matrix Analysis*, Grad. Texts in Math. 169, Springer-Verlag, New York, 1997.

[3] J. C. ENGWERDA, A. C. M. RAN, AND A. L. RIJKEBOER, *Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^\star X^{-1}A = Q$*, Linear Algebra Appl., 186 (1993), pp. 255–275.

[4] J. C. ENGWERDA, *On the existence of the positive definite solution of the matrix equation $X + A^TX^{-1}A = I$*, Linear Algebra Appl., 194 (1993), pp. 91–108.

[5] A. FERRANTE AND B. C. LEVY, *Hermitian solutions of the equation $X = Q + NX^{-1}N^\star$*, Linear Algebra Appl., 247 (1996), pp. 359–373.

[6] C.-H. GUO AND P. LANCASTER, *Iterative solution of two matrix equations*, Math. Comp., 68 (1999), pp. 1589–1603.

[7] H. HEUSER, *Funktionalanalysis*, Teubner, Stuttgart, 1975.

[8] I. G. IVANOV AND S. M. EL-SAYED, *Properties of positive definite solutions of the equation $X + A^\star X^{-2}A = I$*, Linear Algebra Appl., 279 (1998), pp. 303–316.

[9] P. LANCASTER AND M. TISMENETSKY, *The Theory of Matrices*, 2nd ed., Academic Press, Orlando, FL, 1985.

[10] P. LANCASTER AND L. RODMAN, *Algebraic Riccati Equations*, Oxford Science, New York, 1995.

[11] A. C. M. RAN AND R. VREUGDENHIL, *Existence and comparison theorems for algebraic Riccati equations for continuous- and discrete-time systems*, Linear Algebra Appl., 99 (1988), pp. 63–83.

[12] S. M. EL-SAYED, *The Study on Special Matrices and Numerical Methods for Special Matrix Equations*, Ph.D. thesis, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria, 1996.

[13] S. M. EL-SAYED, *Theorems for the existence and computing of positive definite solutions for two nonlinear matrix equations*, in Proceedings of 25th Spring Conference of the Union of Bulgarian Mathematicians, Bulgarian Academy of Sciences, Kazanlak, 1996.

[14] V. SIMA, *Algorithms for Linear-Quadratic Optimization*, Pure Appl. Math. 200, Marcel Dekker, New York, 1996.

[15] W.-Y. YAN, J. B. MOORE, AND U. HELMKE, *Recursive algorithms for solving a class of nonlinear matrix equations with applications to certain sensitivity optimization problems*, SIAM J. Control Optim., 32 (1994), pp. 1559–1576.

[16] X. ZHAN, *Computing the extremal positive definite solutions of a matrix equation*, SIAM J. Sci. Comput., 17 (1996), pp. 1167–1174.

[17] X. ZHAN AND J. XIE, *On the matrix equation $X + A^TX^{-1}A = I$*, Linear Algebra Appl., 247, (1996), pp. 337–345.