ON A NEWTON-LIKE METHOD FOR SOLVING ALGEBRAIC RICCATI EQUATIONS*

CHUN-HUA GUO† AND ALAN J. LAUB‡

Abstract. An exact line search method has been introduced by Benner and Byers [*IEEE Trans. Autom. Control*, 43 (1998), pp. 101–107] for solving continuous algebraic Riccati equations. The method is a modification of Newton's method. A convergence theory is established in that paper for the Newton-like method under the strong hypothesis of controllability, while the original Newton's method needs only the weaker hypothesis of stabilizability for its convergence theory. It is conjectured there that the controllability condition can be weakened to the stabilizability condition. In this note we prove that conjecture.

Key words. continuous algebraic Riccati equations, Newton's method, maximal symmetric solution, convergence rate, line search

AMS subject classifications. 15A24, 65F30, 93B40

1. Introduction. Consider the continuous algebraic Riccati equation (CARE)

$$\mathcal{R}(X) = XDX - XA - A^TX - C = 0, \tag{1.1}$$

where $A, D, C \in \mathbb{R}^{n\times n}$, and $D^T = D$, $C^T = C$. Matrix equations of this form occur in many important applications (see [11], [9]). We limit our attention to real equations, since the complex analog of these equations can be studied in a similar manner. All matrices in this note are real matrices.

For $A, H \in \mathbb{R}^{n\times n}$, the pair (A, H) is said to be controllable if

$$\operatorname{rank}(H \;\, AH \;\, A^2H \;\, \cdots \;\, A^{n-1}H) = n.$$

The pair (A, H) is called stabilizable if there is a $K \in \mathbb{R}^{n\times n}$ such that $A - HK$ is stable, i.e., all its eigenvalues are in the open left half-plane. Note that the matrix K can be chosen to be symmetric if the matrix H is symmetric (see [9]). It is well known that controllability implies stabilizability (see [9]). The order relation on the set of symmetric matrices is the usual one: $X \geq Y$ if $X - Y$ is positive semidefinite. A symmetric solution X_+ of (1.1) is called maximal if $X_+ \geq X$ for every symmetric solution X.

Let $\mathcal{S}$ be the set of symmetric matrices in $\mathbb{R}^{n\times n}$. The Riccati function $\mathcal{R}$ is clearly a mapping from $\mathcal{S}$ into itself. Moreover, the first Fréchet derivative of $\mathcal{R}$ at a matrix X is a linear map $\mathcal{R}'_X : \mathcal{S} \to \mathcal{S}$ given by

$$\mathcal{R}'_X(S) = -\{(A - DX)^TS + S(A - DX)\}.$$

When (A, D) is stabilizable, a Newton procedure for the solution of the Riccati equation (1.1) is then as follows.

ALGORITHM 1.1. Newton's method for the CARE

1. *Choose a symmetric matrix X_0 for which $A - DX_0$ is stable.*

*This research was supported in part by National Science Foundation grant ECS-9633326.

†Department of Computer Science, University of California, Davis, One Shields Avenue, Davis, CA 95616-8562 (guo@cs.ucdavis.edu).

‡College of Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616-5294 (laub@ucdavis.edu).

2. *For $k = 0, 1, \dots$ do:*
   Solve the Lyapunov equation $(A - DX_k)^T N_k + N_k(A - DX_k) = \mathcal{R}(X_k)$;
   Compute $X_{k+1} = X_k + N_k$.

Note that the matrix X_0 in the above algorithm can be found by the method described in [12]. The Lyapunov equation has a unique solution whenever the matrix $A - DX_k$ is stable (see [9], for example). The unique solution N_k can be found efficiently by the methods described in [1].

Concerning the convergence of Algorithm 1.1, we have the following result (see [8], [4], [5], and [9]).

THEOREM 1.2. *Assume that $D \geq 0$, $C^T = C$, (A, D) is stabilizable, and there exists a symmetric solution of the CARE* (1.1)*. Then there is a maximal symmetric solution X_+ of* (1.1) *and all the eigenvalues of $A - DX_+$ are in the closed left half-plane. Moreover, Algorithm* 1.1 *determines a sequence of symmetric matrices $\{X_k\}_{k=1}^{\infty}$ for which $A - DX_k$ is stable for $k = 1, 2, \dots$, $X_1 \geq X_2 \geq \cdots$, and $\lim_{k\to\infty} X_k = X_+$. The convergence is quadratic if $A - DX_+$ has no eigenvalues on the imaginary axis.*

Remark 1.1. The situation where $A - DX_+$ has eigenvalues on the imaginary axis has been studied recently in [6]. It is shown in [6] that the convergence of Algorithm 1.1 is either quadratic or linear with rate 1/2 provided that all eigenvalues of $A - DX_+$ on the imaginary axis are semisimple. Whether quadratic convergence is indeed possible remains an open problem. If the convergence is linear, the performance of Algorithm 1.1 can be improved dramatically by using a simple modification strategy (see [6] for details). When $A - DX_+$ has non-semisimple eigenvalues on the imaginary axis the convergence behavior of Algorithm 1.1 also remains an open problem.

In this note we limit our attention to the case where $A - DX_+$ has no eigenvalues on the imaginary axis. In this case X_+ is called a stabilizing solution.

It is important to note that $X_0 \geq X_1$ is not true in general. As an example (see [7]), consider the CARE (1.1) with $A = 0$, $D = I$, $C = I$. In this case $X_+ = I$, but if $X_0 = \epsilon I$ (ϵ is a small positive number), we have $X_1 = \frac{1+\epsilon^2}{2\epsilon} I$. Thus $\|X_1 - X_+\|_\infty$ can be arbitrarily large even though $\|X_0 - X_+\|_\infty < 1$. In [7], some tests are described to determine whether Newton refinement will improve an approximate solution obtained by other methods. In [3], Benner and Byers introduced step size control into Newton's method for the generalized CARE. The generalized CARE they considered can be reduced to the CARE of the form (1.1). Although the reduction is not always numerically feasible, convergence analysis can always be made on the reduced equation. We can therefore limit our attention to the CARE of the form (1.1).

ALGORITHM 1.3. Exact line search method [3]

1. *Choose a symmetric matrix X_0 for which $A - DX_0$ is stable.*
2. *For $k = 0, 1, \dots$ do:*
   Solve the Lyapunov equation $(A - DX_k)^T N_k + N_k(A - DX_k) = \mathcal{R}(X_k)$;
   Find a minimizer t_k of $f_k(t) = \|\mathcal{R}(X_k + tN_k)\|_F^2$ $(0 \leq t \leq 2)$;
   Compute $X_{k+1} = X_k + t_k N_k$.

In the above algorithm, $\|\cdot\|_F$ is the Frobenius norm. The algorithm effectively solves the problem of a potentially disastrous first Newton step and can be implemented in such a way that the increased amount of computational work per iteration is marginal (see [3] for details).

The following convergence result is proved in [3].

THEOREM 1.4. *Assume that $D \geq 0$, $C^T = C$, (A, D) is controllable, and there is a stabilizing maximal solution X_+ of the CARE* (1.1)*. If the minimizer t_k of*

$f_k(t)$ $(0 \le t \le 2)$ satisfies $t_k \ge t_L > 0$ for all $k \ge 0$, where t_L is a constant tolerance threshold, then the sequence $\{X_k\}$ in Algorithm 1.3 *is well defined and $X_k \to X_+$ quadratically. If $t_k = 0$ for some $k \ge 0$, then $X_k = X_+$.*

In Theorem 1.2, we only need the stabilizability of (A, D) for the convergence of Newton's method. In the above theorem, however, the controllability of (A, D) is required. Since Algorithm 1.3 is designed to improve the performance of Newton's method, it would be desirable to replace the controllability condition in the above theorem by the stabilizability condition. It is conjectured in [3] that this can be done. We will confirm the conjecture in this note.

2. Relaxing the controllability condition. Several interesting results are established in [3] to prove Theorem 1.4. The following two lemmas will be utilized to relax the controllability condition in Theorem 1.4.

LEMMA 2.1 (see [3]). *Assume that $D \ge 0$, $C^T = C$, and $A - DX_+$ is stable. If $A - DX_k$ is stable, then $A - D(X_k + tN_k)$ is also stable for any $t \in [0, 2]$.*

Remark 2.1. The controllability of (A, D) is not needed for this lemma. In the proof (given in [3]) of Theorem 1.4, the controllability of (A, D) is required to prove that the sequence $\{X_k\}$ (already well defined by Lemma 2.1) is bounded. Otherwise the stabilizability of (A, D) is sufficient for the proof of Theorem 1.4.

LEMMA 2.2 (see [3]). *Suppose that $\{X_k\}_{k=1}^{\infty}$ is a sequence of symmetric matrices such that $\{\mathcal{R}(X_k)\}_{k=1}^{\infty}$ is bounded. If (A, D) is controllable, then $\{X_k\}_{k=1}^{\infty}$ is bounded.*

Remark 2.2. In the above lemma, the controllability of (A, D) cannot be replaced by the stabilizability of (A, D). This is illustrated by the CARE (1.1) with

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

(see Example 7.9.1 of [9]). Clearly, the pair (A, D) is stabilizable but not controllable. The real symmetric solutions of the CARE can be found to be

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & b \\ b & -b^2/2 \end{pmatrix}, \quad b \in \mathbb{R}.$$

Note that the solution set of $\mathcal{R}(X) = 0$ is already unbounded. Note however that among these solutions, there is only one stabilizing solution, namely $X_+ = \operatorname{diag}(1, 0)$. The uniqueness of the stabilizing solution is actually a general result (see Theorem 9.3.1 of [9], for example).

The above observation leads us to formulate the following result (see also Acknowledgments).

LEMMA 2.3. *Suppose that $\{X_k\}_{k=1}^{\infty}$ is a sequence of symmetric matrices such that $\{\mathcal{R}(X_k)\}_{k=1}^{\infty}$ is bounded. If (A, D) is stabilizable and $A - DX_k$ is stable for each $k \ge 1$, then $\{X_k\}_{k=1}^{\infty}$ is bounded.*

Proof. It is clear that we may assume $C = 0$ without loss of generality. Consider the controllable subspace of (A, D):

$$\mathcal{C} = \operatorname{Im}(D \; AD \; A^2D \; \cdots \; A^{n-1}D).$$

If $\dim \mathcal{C} = n$, then (A, D) is controllable and the result follows directly from Lemma 2.2. If $\dim \mathcal{C} = r < n$, we can consider using the control (or Kalman) normal form of (A, D) (see [9]). Thus, we take an orthonormal basis for $\mathcal{C}$ and expand it to an orthonormal basis of $\mathbb{R}^n$. Under this basis, the matrices A and D have very simple forms. More precisely, there is an orthogonal matrix U such that

$$A = U^T \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} U, \quad D = U^T \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} U,$$

where $A_1, D_1 \in \mathbb{R}^{r \times r}$. Write accordingly

$$X_k = U^T \begin{pmatrix} X_{k,1} & X_{k,3} \\ X_{k,3}^T & X_{k,2} \end{pmatrix} U.$$

Since $A - DX_k$ is stable for all $k \geq 1$, we see immediately that A_2 is stable and $A_1 - D_1 X_{k,1}$ is stable for all $k \geq 1$. Moreover, it follows readily from $\dim \mathcal{C} = r$ that

$$\dim \operatorname{Im}(D_1 \ A_1 D_1 \ \cdots \ A_1^{r-1} D_1) = \dim \operatorname{Im}(D_1 \ A_1 D_1 \ \cdots \ A_1^{n-1} D_1) = r.$$

Thus, (A_1, D_1) is controllable.

Since the sequence $\{\mathcal{R}(X_k)\}$ is bounded, direct computation shows that the matrix sequences

$$\{X_{k,1} D_1 X_{k,1} - X_{k,1} A_1 - A_1^T X_{k,1}\}, \tag{2.1}$$

$$\{X_{k,1} D_1 X_{k,3} - X_{k,1} A_3 - X_{k,3} A_2 - A_1^T X_{k,3}\}, \tag{2.2}$$

and

$$\{X_{k,3}^T D_1 X_{k,3} - X_{k,3}^T A_3 - X_{k,2} A_2 - A_3^T X_{k,3} - A_2^T X_{k,2}\} \tag{2.3}$$

are all bounded. Since (2.1) is bounded and (A_1, D_1) is controllable, we know from Lemma 2.2 that $\{X_{k,1}\}$ is bounded. We then know from the boundedness of (2.2) that $\{(A_1 - D_1 X_{k,1})^T X_{k,3} + X_{k,3} A_2\}$ is bounded. Since $\{A_1 - D_1 X_{k,1}\}$ is bounded, A_2 is stable, and $A_1 - D_1 X_{k,1}$ is stable for all $k \geq 1$, it follows from Lemma 2.4 below that $\{X_{k,3}\}$ is bounded. Now that $\{X_{k,3}\}$ has been shown to be bounded, it follows from the boundedness of (2.3) that $\{X_{k,2} A_2 + A_2^T X_{k,2}\}$ is bounded. Since A_2 is stable, $\{X_{k,2}\}$ is also bounded by Lemma 2.4. Therefore, $\{X_k\}$ is bounded. □

The following lemma has been used in the proof of Lemma 2.3.

LEMMA 2.4. *Let $\mathcal{A}, \mathcal{B}$, and $\mathcal{C}$ be bounded sets of matrices of size $m \times m$, $n \times n$, and $m \times n$, respectively. Assume that the closure of $\bigcup_{A \in \mathcal{A}} \sigma(A)$ and the closure of $\bigcup_{B \in \mathcal{B}} \sigma(B)$ are disjoint. Then the set*

$$\mathcal{G} = \{X \mid AX - XB = C,\ A \in \mathcal{A},\ B \in \mathcal{B},\ C \in \mathcal{C}\}$$

is bounded.

Proof. For $A \in \operatorname{clo}(\mathcal{A})$, $B \in \operatorname{clo}(\mathcal{B})$, and $C \in \operatorname{clo}(\mathcal{C})$, the Sylvester equation $AX - XB = C$ has a unique solution $X = F(A, B, C)$ by the assumption. See, for example, [9] or [10]. It is clear that the function F is continuous on the compact set $\operatorname{clo}(\mathcal{A}) \times \operatorname{clo}(\mathcal{B}) \times \operatorname{clo}(\mathcal{C})$. Therefore F maps this compact set onto a compact set. As a result, $\mathcal{G}$ is bounded. □

We can now relax the controllability condition in Theorem 1.4.

THEOREM 2.5. *The conclusions in Theorem* 1.4 *are valid if the pair (A, D) is stabilizable.*

Proof. This follows readily from Remark 2.1, Lemma 2.1, and Lemma 2.3. Note that $\{\mathcal{R}(X_k)\}_{k=1}^{\infty}$ is bounded by Algorithm 1.3 itself. □

Acknowledgments. The authors thank Peter Benner for pointing out to them that Lemma 2.3 was formulated earlier as a conjecture in [2] and for sending them a copy of the reference. Chun-Hua Guo would like to thank Peter Lancaster for introducing him to the study of algebraic Riccati equations several years ago. He also gratefully acknowledges the support of an NSERC postdoctoral fellowship.

REFERENCES

[1] R. H. Bartels and G. W. Stewart, *Solution of the matrix equation $AX + XB = C$*, Comm. ACM, 15 (1972), pp. 820–826.
[2] P. Benner, *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems*, Logos-Verlag, Berlin, 1997.
[3] P. Benner and R. Byers, *An exact line search method for solving generalized continuous-time algebraic Riccati equations*, IEEE Trans. Autom. Control, 43 (1998), pp. 101–107.
[4] W. A. Coppel, *Matrix quadratic equations*, Bull. Austral. Math. Soc., 10 (1974), pp. 377–401.
[5] I. Gohberg, P. Lancaster, and L. Rodman, *On Hermitian solutions of the symmetric algebraic Riccati equation*, SIAM J. Control Optimization, 24 (1986), pp. 1323–1334.
[6] C.-H. Guo and P. Lancaster, *Analysis and modification of Newton's method for algebraic Riccati equations*, Math. Comp., 67 (1998), pp. 1089–1105.
[7] C. Kenney, A. J. Laub, and M. Wette, *Error bounds for Newton refinement of solutions to algebraic Riccati equations*, Math. Control Signals Systems, 3 (1990), pp. 211–224.
[8] D. L. Kleinman, *On an iterative technique for Riccati equation computations*, IEEE Trans. Autom. Control, 13 (1968), pp. 114–115
[9] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.
[10] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Second Edition, Academic Press, Orlando, 1985.
[11] V. L. Mehrmann, *The Autonomous Linear Quadratic Control Problem*, Lecture Notes in Control and Information Sciences, Vol. 163, Springer-Verlag, Berlin, 1991.
[12] V. Sima, *An efficient Schur method to solve the stabilizing problem*, IEEE Trans. Autom. Control, 26 (1981), pp. 724–725.