Comparison Theorems for the Matrix Riccati Equation

Jonq Juang* and Mei Tzu Lee
Department of Applied Mathematics
National Chiao Tung University
Hsinchu, Taiwan, R.O.C.

Submitted by Richard A. Brualdi

ABSTRACT

We derive comparison theorems for the matrix-valued Riccati equations of the form $R_i'(z) = B_i(z) + A_i(z)R_i(z) + R_i(z)D_i(z) + R_i(z)C_i(z)R_i(z)$, $R_i(0) = R_{0,i}$, $i = 1, 2$. Such equations arise in transport problems and other applications. Sufficient conditions under which $R_1(z) - R_2(z) \geqslant 0$ and $R_1(z) - R_2(z) > 0$ (in the componentwise sense) for all z are given, respectively. The comparison theorems in which the positivity is defined in terms of positive semidefiniteness were obtained by Royden. Because of the intrinsic disparity in the nature of positivity structure, the techniques developed there cannot be applied to our problem.

1. INTRODUCTION

Consider the matrix-valued Riccati equations of the form

$$\begin{aligned} R'(z) &= B(z) + A(z)R(z) + R(z)D(z) + R(z)C(z)R(z), \\ R(0) &= R_0. \end{aligned} \tag{1}$$

We assume that coefficient matrices A, B, C, and D of appropriate dimensions are piecewise continuous on $[0, a)$.

* The work was partially supported by the National Science Council of the Republic of China.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

Such equations arise in transport theory (see e.g., [1–4, 6]) and other applications. We refer to [1], [2], and [6] for further physical background regarding this problem.

Let the matrix $E(z)$ be defined in block form by

$$E(z) = \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix}; \tag{2}$$

it is called the matrix of the Riccati equation. If G and H are two real matrices, we write $G \geqslant H$ to denote that $G_{ij} - H_{ij} \geqslant 0$, for all i, j, i.e., the element in the ith row and jth column of G is no less than that of H. If $G_{ij} - H_{ij} > 0$ for all i, j, then we write $G > H$.

In Section 2 we introduce notation and underlying assumptions, and record some well-known results for completeness and ease of reference. Section 4 is devoted to the main results. Specifically, if R_1 satisfies a Riccati equation of the form (1) with matrix $E_1(z)$ and if R_2 satisfies an equation of the same form with matrix $E_2(z)$, then we establish conditions under which $R_1(z) \geqslant R_2(z)$ for all z and $R_1(z) > R_2(z)$ for all z, respectively. The remainder of this introductory section consists of an overview of previous work related to the present work.

The earlier study most closely related to ours are that of Royden [5]. In [5], the positivity structure is defined in the sense of positive semidefiniteness. Other differences are that in [5] the coefficient matrices are independent of z, and that the matrices B and C are required to be symmetric and $A = D^T$. Perhaps most importantly, because the difference in the positivity structure, the techniques developed in [5] in obtaining comparison theorems cannot be applied to our problem. A different approach needs to be developed.

2. PRELIMINARIES

In the following discussion the matrix functions A, B, C, and D will be supposed to be defined on a given interval $[0, \infty)$ on the real line, and satisfy the following hypothesis:

> $A(z)$, $B(z)$, $C(z)$, and $D(z)$ are matrices of appropriate dimensions, which on arbitrary compact subintervals $[a, b]$ of $[0, \infty)$ are, respectively, piecewise continuous. (3)

NOTATION. Let x_E be such that $[0, x_E)$ is the maximal interval of existence of a solution.

It is clear that under this hypothesis the equation (1) with the matrix $E(z)$ has a unique solution on the maximal interval $[0, x_E)$ of existence. We furthermore assume that the coefficient matrices satisfy the additional conditions

$$B(z) \geqslant 0, \qquad C(z) \geqslant 0 \qquad \text{for } z \text{ a.e. on } [0, x_E), \tag{5a}$$

$$A(z) {}^* \geqslant {}^* 0, \qquad D(z) {}^* \geqslant {}^* 0 \qquad \text{for } z \text{ a.e. on } [0, x_E). \tag{5b}$$

Here (5b) denotes that $A(z)$ and $D(z)$ have nonnegative entries except possibly along their diagonals. Except for a change of notation the following result is essentially Theorem 9.2 of Chapter 2 of Reid [4], and we refer to this source for a proof.

THEOREM 1. *Suppose the matrix $E(z)$ of* (2) *satisfies* (3) *and* (5) *and the initial value $R_0 \geqslant 0$. Then the solution $R(z)$ is unique and has nonnegative entries for $z \in [0, x_E)$.*

In what follows, we shall introduce the concept of irreducibility (see e.g. [7]).

DEFINITION 1. A matrix $A = (a_{ij})$ of order N is irreducible if $N = 1$ or if $N > 1$ and given any two nonempty disjoint subsets S and T of $W := \{1, 2, \dots, N\}$, and whose union is W, there exist $i \in S$ and $j \in T$ such that $a_{ij} \neq 0$.

An equivalent statement of irreducibility is given in the following:

THEOREM 2 (See e.g. Theorem 2.5.2 of [7]). *A matrix A of order $N > 1$ is irreducible if and only if given any two distinct integers i and j, $1 \leqslant i, j \leqslant N$, either $a_{ij} \neq 0$ or there exist $i_1, i_2, \dots, i_s$ such that $a_{ii_1} a_{i_1 i_2} \cdots a_{i_s j} \neq 0$.*

3. A COMPARISON THEOREM FOR MATRIX RICCATI EQUATIONS

Our object in this section is to derive some comparison theorems for matrix Riccati equations of the form (1). We shall assume throughout that the matrices $E_i(z)$, $i = 1, 2$, of Equation (1) satisfy (3) and (5), and that the initial values R_{01}, $R_{02} \geqslant 0$.

THEOREM 3. *Let R_1 and R_2 be, respectively, solutions of Equation* (1) *with matrices $E_1(z)$ and $E_2(z)$, and with initial values R_{01} and R_{02}. Suppose that*

$$E_1(z) \leqslant E_2(z) \tag{6}$$

and that

$$R_{01} \leqslant R_{02}. \tag{7}$$

Then $R_1(z) \leqslant R_2(z)$ on $[0, b)$, where $b = \min\{x_{E_1}, x_{E_2}\}$. If, in addition, $R_{01} < R_{02}$, then $R_1(z) < R_2(z)$.

Proof. Define $\tilde{A}_i(z) = A_i(z) - \operatorname{diag} A_i(z)$, $\tilde{D}_i(z) = D_i(z) - \operatorname{diag} D_i(z)$, $i = 1, 2$. We have immediately, for $i = 1, 2$, that

$$\begin{aligned} &R_i'(s) - [\operatorname{diag} A_i(s)] R_i(s) - R_i(s)[\operatorname{diag} D_i(s)] \\ &\qquad = B_i(s) + \tilde{A}_i(s) R_i(s) + R_i(s) \tilde{D}_i(s) + R_i(s) C_i(s) R_i(s). \end{aligned} \tag{8}$$

Premultiplying and postmultiplying the equation (8) by the integration factors $\exp \int_s^z \operatorname{diag} A_i(t)\, dt$ and $\exp \int_s^z \operatorname{diag} D_i(t)\, dt$, respectively, and integrating the resulting equation with respect to s from 0 to z, we obtain, for $i = 1, 2$, that

$$\begin{aligned} R_i(z) &= \exp\left(\int_0^z \operatorname{diag} A_i(t)\, dt\right) R_{0i} \exp\left(\int_0^z \operatorname{diag} D_i(t)\, dt\right) \\ &\quad + \int_0^z \exp\left(\int_s^z \operatorname{diag} A_i(t)\, dt\right) \\ &\qquad \times \left[B_i(s) + \tilde{A}_i(s) R_i(s) + R_i(s) \tilde{D}_i(s) + R_i(s) C_i(s) R_i(s) \right] \\ &\qquad \times \exp\left(\int_s^z \operatorname{diag} D_i(t)\, dt\right) ds. \\ &:= \mathcal{G}_i R_i(z). \end{aligned} \tag{9}$$

Let us define the standard Picard iteration $\{R^{(m)}(z)\}$ by

$$\begin{aligned} R_i^{(0)}(z) &= \exp\left(\int_0^z \operatorname{diag} A_i(t)\,dt\right) R_{0i} \exp\left(\int_0^z \operatorname{diag} D_i(t)\,dt\right), \\ R_i^{(m+1)}(z) &= \mathcal{G}_i R_i^{(m)}(z), \end{aligned} \tag{10}$$

where $\mathcal{G}_i$ is defined in (9). It is clear, via the assumptions (3) and (5), that $\mathcal{G}_i$ is a monotone operator. An easy induction gives that

$$0 \leqslant R_i^{(m)}(z) \leqslant R_i^{(m+1)}(z) \leqslant R_i(z) \tag{11}$$

for every $z \in [0, b)$ and each $m = 0, 1, \ldots$. Moreover, it follows readily from the assumptions (6) and (7) that

$$0 \leqslant R_1^{(m)}(z) \leqslant R_2^{(m)}(z) \tag{12}$$

for every $z \in [0, b)$ and each $m = 0, 1, \ldots$. From (11) and (12), we have that $R_1(z) \leqslant R_2(z)\ \forall z \in [0, b)$.

The last assertion of the theorem follows from (3), (5), (6), (7), and (9). This completes the proof of Theorem 3. ∎

The remainder of this section is devoted to the problem of finding nontrivial sufficient conditions on $E_i(z)$ so that one solution is strictly greater than the other.

Let the matrix $M(z)$ be defined by

$$M(z) = \begin{pmatrix} \tilde{A}_2(z) - \tilde{A}_1(z) & B_2(z) - B_1(z) \\ C_2(z) - C_1(z) & \tilde{D}_2(z) - \tilde{D}_1(z) \end{pmatrix} = \left(m_{ij}(z)\right). \tag{13}$$

For simplicity, we will assume here that the matrices $A_i(z)$, $B_i(z)$, $C_i(z)$, and $D_i(z)$, $i = 1, 2$ are $n \times n$. For each $z \geqslant 0$, let $G(z)$ be the directed graph on the vertices $1, 2, \ldots, 2n$ such that the edge $(i, j) \in G(z)$ if and only if $m_{ij} > 0$.

THEOREM 4. *Suppose the hypothesis of Theorem 3 holds. Suppose, furthermore, there exists an $\varepsilon > 0$ such that for every pair of indices $i, n+j$, $1 \leqslant i, j \leqslant n$, there is a directed path from i to $n+j$ in $G(z)$ for all z in $(0, \varepsilon)$. Then $R_2(z) > R_1(z)$ for $z \in (0, b)$.*

Proof. Using (9), we obtain that, for $z \in (0, b)$,

$$\begin{aligned}
R_2(z) - R_1(z) = {} & \left[R_2^{(0)}(z) - R_1^{(0)}(z)\right] + \int_0^z \exp\left(\int_s^z \operatorname{diag} A_2(t)\,dt\right) \\
& \times\left[B_2(s) + \tilde{A}_2(s)R_2(s) + R_2(s)\tilde{D}_2(s) \right. \\
& \left. + R_2(s)C_2(s)R_2(s)\right] \\
& \times \exp\left(\int_s^z \operatorname{diag} D_2(t)\,dt\right) ds \\
& - \int_0^z \exp\left(\int_s^z \operatorname{diag} A_1(t)\,dt\right) \\
& \times\left[B_1(s) + \tilde{A}_1(s)R_1(s) + R_1(s)\tilde{D}_1(s) \right. \\
& \left. + R_1(s)C_1(s)R_1(s)\right] \\
& \times \exp\left(\int_s^z \operatorname{diag} D_1(t)\,dt\right) ds.
\end{aligned}$$

By the facts that $\operatorname{diag} A_1(z) \leqslant \operatorname{diag} A_2(z)$, $\operatorname{diag} D_1(z) \leqslant \operatorname{diag} D_2(z)$, and $R_{01}(z) \leqslant R_{02}(z)$ for $z \in [0, b)$, we have that

$$\begin{aligned}
R_2(z) - R_1(z) \geqslant {} & \int_0^z \exp\left(\int_s^z \operatorname{diag} A_1(t)\,dt\right)\Big\{\left[B_2(s) - B_1(s)\right] \\
& + \left[\tilde{A}_2(s)R_2(s) - \tilde{A}_1(s)R_1(s)\right] \\
& + \left[R_2(s)\tilde{D}_2(s) - R_1(s)\tilde{D}_1(s)\right] \\
& + \left[R_2(s)C_2(s)R_2(s) - R_1(s)C_1(s)R_1(z)\right]\Big\} \\
& \times \exp\left(\int_s^z \operatorname{diag} D_1(t)\,dt\right) ds.
\end{aligned}$$

To further estimate $R_2(z) - R_1(z)$, we have, after some calculations, that, for $z \in [0, b)$,

$$R_2(z)C_2(z)R_2(z) - R_1(z)C_1(z)R_1(z)$$
$$\geqslant [R_2(z) - R_1(z)][C_2(z)R_2(z) - C_1(z)R_1(z)]$$
$$\geqslant [R_2(z) - R_1(z)][C_2(z) - C_1(z)][R_2(z) - R_1(z)].$$

The first inequality above is justified by Theorem 3, the assumption, and the fact that

$$R_1(z)C_2(z)R_2(z) + R_2(z)C_1(z)R_1(z) \geqslant 2R_1(z)C_1(z)R_1(z).$$

Similarly, we obtain

$$R_2(z)\tilde{D}_2(z) - R_1(z)\tilde{D}_1(z) \geqslant [R_2(z) - R_1(z)]\left[\tilde{D}_2(z) - \tilde{D}_1(z)\right],$$

and

$$\tilde{A}_2(z)R_2(z) - \tilde{A}_1(z)R_1(z) \geqslant \left[\tilde{A}_2(z) - \tilde{A}_1(z)\right][R_2(z) - R_1(z)].$$

Using the above estimates, we have that

$$\begin{aligned} R_2(z) - R_1(z) \geqslant \int_0^z &\exp\left(\int_s^z \operatorname{diag} A_1(t)\, dt\right)\Big\{[B_2(s) - B_1(s)] \\ &+ \left[\tilde{A}_2(s) - \tilde{A}_1(s)\right][R_2(s) - R_1(s)] \\ &+ [R_2(s) - R_1(s)]\left[\tilde{D}_2(s) - \tilde{D}_1(s)\right] \\ &+ [R_2(s) - R_1(s)][C_2(s) - C_1(s)][R_2(s) - R_1(s)]\Big\} \\ &\times \exp\left(\int_0^z \operatorname{diag} D_1(t)\, dt\right) dt. \end{aligned} \tag{14}$$

To complete the proof, it suffices to show that the right-hand side of (13) is greater than 0. To this end, let $P = i_0, i_1, \ldots, i_m$ be a sequence of $m + 1$ distinct vertices between 1 and $2n$ such that $i_0 = i \in U$, $i_m = n + j \in T$, where $U = \{1, 2, \ldots, n\}$, $T = \{n + 1, n + 2, \ldots, 2n\}$. We now proceed to show that $(R_2(z) - R_1(z))_{ij} > 0$ for $1 \leqslant i, j \leqslant n$, by induction on m, the length of the path P. For each i, j, if $m = 1$, then, via (14),

$$\begin{aligned}(R_2(z) - R_1(z))_{ij} \geqslant \int_{[0,z]\cap(0,\varepsilon)} &\exp\left(\int_s^z (\operatorname{diag} A_1(t))_{ii}\,dt\right)\\ &\times (B_2(s)) - B_1(s))_{ij} \exp\left(\int_s^z (\operatorname{diag} D_1(t))_{jj}\,dt\right) ds\\ > 0.\end{aligned}$$

Suppose the desired result is true for paths of length less than or equal to k, and let $P = i_0, i_1, \ldots, i_{k+1}$, be a fixed path of length $k + 1$ with $i_0 = i \in U$, $i_{k+1} = n + j \in T$. If $i_k = n + l \in T$, then $(R_2(z) - R_1(z))_{il} > 0$ and $(\tilde{D}_2(z) - \tilde{D}_1(z))_{lj} > 0$ for $z \in [0, b)$. Thus,

$$\begin{aligned}(R_2(z) - R_1(z))_{ij} \geqslant \int_{(0,\varepsilon)\cap[0,z)} &\exp\left(\int_s^z (\operatorname{diag} A_1(t))_{ii}\,dt\right)\\ &\times (R_2(z) - R_1(z))_{il}(\tilde{D}_2(z) - \tilde{D}_1(z))_{lj}\\ &\times \exp\left(\int_s^z (\operatorname{diag} D_1(t))_{jj}\,dt\right) ds > 0.\end{aligned}$$

If $i_1 \in U$, the same conclusion follows from a similar estimate applied to the term $[\tilde{A}_2(s) - \tilde{A}_1(s)][R_2(s) - R_1(s)]$ occurring on the right side of (13). Finally, if neither $i_k \in T$ nor $i_1 \in U$, then necessarily there exists p, $1 \leqslant p \leqslant k - 1$, such that $i_p = q + n \in T$ and $i_{p+1} = r \in U$. It follows from the inductive assumption that both $(R_2(z) - R_1(z))_{iq}$ and $(R_2(z) - R_1(z))_{rj}$ are positive on $[0, b)$. The desired conclusion then comes from the estimate

$$\begin{aligned}(R_2(z) - R_1(z))_{ij} \geqslant \int_{(0,\varepsilon)\cap[0,z]} &\exp\left(\int_s^z (\operatorname{diag} A(t))_{ii}\,dt\right)\\ &\times (R_2(s) - R_1(s))_{iq}(C_2(s) - C_1(s))_{qr}\\ &\times (R_2(s) - R_1(s))_{rj}\\ &\times \exp\left(\int_s^z (\operatorname{diag} D(t))_{jj}\,dt\right) ds > 0.\end{aligned}$$

This completes the proof of Theorem 4. ∎

The following corollary is a direct consequence of Theorems 2 and 4.

COROLLARY 1. *Suppose* $M(z)$, *as defined in* (13), *is irreducible for all* z *in* $(0, \varepsilon)$, *for some* $\varepsilon > 0$. *Then the assertion of Theorem* 4 *holds.*

REFERENCES

1 R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York, 1970.
2 R. Bellman and G. M. Wing, *An Introduction to Invariant Imbedding*, Wiley, New York, 1975.
3 G. H. Meyer, *Initial Value Methods for Boundary Value Problems*, Academic, New York, 1973.
4 W. T. Reid, *Riccati Differential Equations*, Academic, New York, 1972.
5 H. L. Royden, Comparison theorems for the matrix Riccati equation, *Comm. Pure Appl. Math.* XLI:739–746 (1988).
6 G. M. Wing, *An Introduction to Transport Theory*, Wiley, New York, 1962.
7 D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, New York, 1971.

Received 25 January 1991; final manuscript accepted 11 January 1993