Matrix quadratic equations

W.A. Coppel

Matrix quadratic equations have found the most diverse applications. The present article gives a connected account of their theory, and contains some new results and new proofs of known results.

1. Introduction

An autonomous linear hamiltonian system of differential equations has the form

$$x' = J^{-1} H x,$$

where $x \in \mathbb{R}^{2n}$ and

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad H = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix}$$

are constant $2n \times 2n$ real matrices which are respectively skew-symmetric, symmetric.\(^1\) It is well-known, and easily established, that there is a close connection between the solutions of (1) and the solutions of the $n \times n$ matrix Riccati equation

$$R[W] \equiv W' + A + WB + B^* W + WCW = 0.$$

The constant solutions of (2) are just the solutions of the matrix quadratic equation

$$Q[W] \equiv A + WB + B^* W + WCW = 0.$$

Largely as a result of this connection, the matrix quadratic equation (3) has found applications in many different fields; for example, optimal

---

Received 4 February 1974.

\(^1\) The conjugate transpose of a matrix is denoted by an asterisk.
control, stability theory, filtering theory, network theory and
differential games. In these applications it is the symmetric solutions of
(3) which are of interest. The object of the present article is to give a
connected account of the theory of the equation (3). References to
previous work are made at various places in the text.

The following identities are easily verified. If $W_1$ and $W_2$ are
solutions of (3) then their difference $D = W_2 - W_1$ satisfies

$$D(B + CW_1) + (B^* + W_1 C)D = 0,$$

(4)

$$D(B + CW_1) + (B^* + W_1 C)D + DCD = 0.$$

Conversely, if $W_1$ is a solution of (3) and if $D$ satisfies (4) or (5),
then $W_2$ is also a solution of (3).

The coefficient matrix

$$M = S^{-1}H = \begin{bmatrix} B & C \\ -A & -B^* \end{bmatrix}$$

has the property

$$JM = -M^*J.$$

(7)

It follows that the characteristic polynomial

$$\varphi(s) = \det(sI-M)$$

is an even function of $s$. If $W$ is a solution of (3) then

$$\varphi(s) = \det(sI-B-CW)\det(sI+B^*+WC),$$

since

$$M \begin{bmatrix} I \\ W \\ 0 \end{bmatrix} = \begin{bmatrix} I \\ W \\ 0 \end{bmatrix} \begin{bmatrix} B+CW & C \\ 0 & -B^*-WC \end{bmatrix}.$$

Finally we recall some standard definitions. A matrix is said to be
stable if all its eigenvalues have negative real part. The inequality
$W_2 \geq W_1$ between two symmetric matrices $W_1$, $W_2$ means that $W_2 - W_1$ is
non-negative definite. If $B$ is an $n \times n$ matrix and $C$ an $n \times m$
matrix the ordered pair $(B, C)$ is said to be controllable if the $n \times mn$
block matrix $[C \ B C \ ... \ B^{n-1} C]$ has rank $n$. For further discussion of this concept see, for example, Kalman et al [3]. If the pair $(B, C)$ is controllable then for any $m \times n$ matrix $W$ the pair $(B + CW, C)$ is also controllable. If $C = C^* \leq 0$ then $(B, C)$ is controllable if and only if

$$\Omega \equiv - \int_0^T e^{-tB}Ce^{-tB^*}dt > 0$$

for some, and hence every, $T > 0$.

**Lemma 1.** If $(B, C)$ is controllable and $C \leq 0$ then $\hat{B} = B + C\Omega^{-1}$ is stable.

Proof. We have

$$B\Omega + \Omega B^* = \int_0^T \left[ \frac{d}{dt} e^{-tB}Ce^{-tB^*} \right]dt$$

$$= e^{-TB}Ce^{-TB^*} - C,$$

and hence

$$\hat{B}\Omega + \hat{B}^* = e^{-TB}Ce^{-TB^*} + C.$$

Let $\lambda$ be an eigenvalue of $\hat{B}^*$ and $\zeta$ a corresponding eigenvector. Then

$$(\lambda + \bar{\lambda})\zeta^*\Omega\zeta = \zeta^*(\hat{B}\Omega + \hat{B}^*)\zeta$$

$$= \zeta^*(e^{-TB}Ce^{-TB^*} + C)\zeta$$

$$\leq 0.$$

Since $\Omega > 0$ it follows that $\Re \lambda \leq 0$. Moreover, if equality holds then $C\zeta = 0$. From the definition of $\hat{B}$ this implies $B^*\zeta = \lambda \zeta$. Hence

$$\Omega\zeta = - \int_0^T e^{-tB}C\zeta e^{-\lambda t}dt = 0.$$

Since $\zeta \neq 0$, this is a contradiction. Hence $\hat{B}$ is stable.

A somewhat different proof of Lemma 1 is given by Lukes [7].

2. Extreme solutions

Throughout this section it will be assumed that $C \leq 0$ and that $(B, C)$ is controllable. Moreover by a solution of (3) we will always mean a symmetric solution, and we will assume that (3) has at least one solution
THEOREM 1. The matrix quadratic equation (3) has maximal and minimal solutions \( W_+ \) and \( W_- \); that is, every solution \( \tilde{W} \) satisfies
\[
W_- \leq \tilde{W} \leq W_+.
\]
All eigenvalues of \( B + CW_+ \) have non-positive real part and all eigenvalues of \( B + CW_- \) have non-negative real part.

We give two proofs. The first uses the relation with the Riccati equation (2). Let \( \tilde{W} \) be a solution of (3) and put
\[
\tilde{B} = B + CW,
\]
\[
U(t) = \int_0^t e^{t\tilde{B}} C e^{\tilde{B}^*}_r dt.
\]
Since \( C \leq 0 \) and \( (\tilde{B}, C) \) is controllable, the symmetric matrix \( U(t) \) is a strictly decreasing function of \( t \). In particular, \( U(t) > 0 \) for \( t < 0 \). As \( t \to -\infty \), \( U^{-1}(t) \) decreases and is bounded below by \( 0 \). Thus \( \lim_{t \to -\infty} U^{-1}(t) \) exists and is non-negative.

The function \( U(t) \) is the solution of the linear differential equation
\[
U' = \tilde{B}U + UB^* + C
\]
which vanishes at \( t = 0 \). For any \( T > 0 \) put
\[
W_T(t) = \tilde{W} + U^{-1}(t-T).
\]
Using (11), it is easily verified that \( W_T(t) \) is a solution of (2) on the interval \([0, T]\). The representation
\[
W_T(t) = [\tilde{W}U(t-T)+I]U^{-1}(t-T)
\]
shows that \( W_T(t) \) is non-singular on a small interval \([T-\delta, T]\) and \( W_T^{-1}(t) \to 0 \) as \( t \to T \). If \( V_T(t) \) is the solution of the 'inverse' Riccati equation...
Matrix quadratic equations

\[ V' - C - VB^* - BV - VAV = 0 \]

such that \( V_T(T) = 0 \) it follows that \( V_T(t) \) is non-singular and

\[ V_T^{-1}(t) = W_T(t) \text{ on } [T-\delta, T). \]

This shows that \( W_T(t) \) can also be defined

independently of \( \tilde{W} \). The existence of \( \tilde{W} \) simply guarantees the existence

of \( W_T(t) \) over the interval \([0, T)\). As \( T \to \infty \), \( W_T(t) \) converges

uniformly on compact \( t \)-intervals to the constant matrix

\[ \tilde{W} = \tilde{W} + \lim_{t \to \infty} U^{-1}(t). \]

Thus \( \tilde{W} \) is a constant solution of the Riccati equation (2); that is, it

is a solution of the quadratic equation (3). Moreover \( \tilde{W} \geq \tilde{W} \) for any

solution \( \hat{W} \) of (3). Since \( (\tilde{B}, \tilde{C}) \) is controllable it follows from Lemma

1 that

\[ B + CW_T(0) = \tilde{B} - C \left[ \int_0^T e^{-\tau \tilde{B}} Ce^{-\tau \tilde{B}^*} \, d\tau \right]^{-1} \]

is a stable matrix. Hence, letting \( T \to \infty \), all eigenvalues of \( B + CW_+ \)

have non-positive real part.

By replacing \( \tilde{W} \) by \( -\tilde{W} \) we see that (3) also has the minimal solution

\[ W = \tilde{W} + \lim_{t \to \infty} U^{-1}(t), \]

and all eigenvalues of \( B + CW_- \) have non-negative real part.

The second proof provides a practical algorithm for the determination

of the maximal and minimal solutions. It is based on the identity

\[ \tilde{W}(B + CW) + (B + CW)^* \tilde{W} - WCW = \tilde{W}(B + CW) + (B + CW)^* \tilde{W} - \tilde{W}C \tilde{W} + (\hat{W} - \tilde{W}) \tilde{C}(\hat{W} - \tilde{W}), \]

valid for arbitrary symmetric matrices \( C, \tilde{W}, \hat{W} \), and on the following simple

**Lemma 2.** Suppose \( C \leq 0 \) and the linear equation

\[ WB + B^*W = C \]

has a solution \( W \geq 0 \). If \( B \) has an eigenvalue \( \lambda = \mu + iv \) with \( \mu \geq 0 \)

and if \( \xi \) is a corresponding eigenvector then \( C\xi = 0 \).
In fact from (15) we obtain at once
\[ \zeta^4 C \zeta = 2 \mu \zeta^4 W \zeta. \]
Since the left side is non-positive and the right side is non-negative they must both be zero. Since \( C \leq 0 \) this implies \( C \zeta = 0 \).

We now continue with the second proof of Theorem 1. By Lemma 1 there exists a symmetric matrix \( H_0 \) such that \( B + CW_0 \) is stable. We define a sequence \( \{W_v\} \) of symmetric matrices inductively by the linear equations
\[
W_{v+1}(B+CW_v) + (B+CW_v)^*W_{v+1} = W_v CW_v - A \quad (v \geq 0).
\]
Suppose \( W_v \) has been defined and \( B + CW_v \) is stable. Then (16) has a unique and symmetric solution \( W_{v+1} \). By the identity (14), with \( W = \tilde{W} \) and \( \tilde{W} = W_v \),
\[
-A = \tilde{W}(B+CW_v) + (B+CW_v)^*\tilde{W} - W_v CW_v + (W_v - \tilde{W})C(W_v - \tilde{W}).
\]
Adding this to (16) we get
\[
(W_{v+1} - \tilde{W})(B+CW_v) + (B+CW_v)^*(W_{v+1} - \tilde{W}) = (W_v - \tilde{W})C(W_v - \tilde{W}).
\]
Therefore
\[
W_{v+1} - \tilde{W} = -\int_0^\infty e^{-t(B+CW_v)^*} (W_v - \tilde{W})C(W_v - \tilde{W}) e^{t(B+CW_v)} dt \\
\geq 0.
\]
By the identity (14) also, with \( W = W_{v+1} \) and \( \tilde{W} = W_v \),
\[
(W_{v+1} + B + CW_{v+1}) + (B + CW_{v+1})^*W_{v+1} - W_{v+1} CW_{v+1} = -A + (W_{v+1} - W_v)C(W_{v+1} - W_v).
\]
Adding (17), with \( v \) replaced by \( v + 1 \), we get
\[
(W_{v+1} - \tilde{W})(B+CW_{v+1}) + (B+CW_{v+1})^*(W_{v+1} - \tilde{W}) = (W_{v+1} - W_v)C(W_{v+1} - W_v) + (\tilde{W} - W_{v+1})C(\tilde{W} - W_{v+1}).
\]
Suppose \( B + CW_{v+1} \) had an eigenvalue \( \lambda \) with non-negative real part, and let \( \zeta \) be a corresponding eigenvector. Then, by Lemma 2,
Matrix quadratic equations

\[ C(W_{v+1} - W_v) \zeta = C(\tilde{W}_{v+1} - \tilde{W}_v) \zeta = 0. \]

Therefore

\[ (B + CW_v) \zeta = (B + CW_{v+1}) \zeta = \lambda \zeta. \]

Since \( B + CW_v \) is stable, this is a contradiction. Thus \( B + CW_{v+1} \) is
stable.

Subtracting (16) from (19) with \( v \) replaced by \( v - 1 \) we get

\[ (W_v - W_{v+1})(B + CW_v) + (B + CW_v)\ast(W_v - W_{v+1}) = (W_{v-1} - W_v)C(W_{v-1} - W_v) \quad (v \geq 1). \]

Therefore

\[
W_v - W_{v+1} = \int_0^\infty e^{t(B + CW_v)} (W_{v-1} - W_v) e^{t(B + CW_v)} dt \\
\geq 0.
\]

Thus, after the first term, the sequence \( \{W_v\} \) is monotonic decreasing and
bounded below by \( \tilde{W} \). It therefore converges to a limit \( \tilde{W}_+ \). Letting
\( v \to \infty \) in (16) we see that \( \tilde{W}_+ \) is a solution of (3). Since \( \tilde{W}_+ \geq \tilde{W} \) for
any solution \( \tilde{W} \) of (3) it is in fact a maximal solution. Since
\( B + CW_+ = \lim(B + CW_v) \), all eigenvalues of \( B + CW_+ \) have non-positive real
part.

The existence of extreme solutions seems to have been first
established by Reid [17], to whom the formulae (12), (13) are also due
(Reid [16]). The algorithm used in the second proof was introduced in a
special case by Kleinman [4] and further developed by Wonham [16]. Its
applicability in the general case is mentioned by Willems [15].

Our next result relates two arbitrary solutions of (3).

**THEOREM 2.** Let \( W_1, W_2 \) be any two solutions of (3) and set

\[ D = W_2 - W_1, \quad B_1 = B + CW_1, \quad B_2 = B + CW_2. \]

Let \( M \) denote the nullspace of \( D \) and let \(^2 m, p, q \) be the dimensions of

\[ m, p, \text{ and } q \text{ are the number of zero, positive, and negative eigenvalues of } D. \]
the largest subspaces invariant under $D$ on which $D = 0$, $D > 0$, $D < 0$.

Then $N$ is invariant under $B_1$ and $B_1 = B_2$ on $N$. Let

$\lambda_1, \ldots, \lambda_m$ be the eigenvalues of the restriction of $B_1$ to $N$ and let

$\lambda_{m+1}, \ldots, \lambda_n$ be the remaining eigenvalues of $B_1$. Then the eigenvalues

of $B_2$ are $\lambda_1, \ldots, \lambda_m, -\lambda_{m+1}, \ldots, -\lambda_n$. Moreover among the $p + q$

eigenvalues $\lambda_{m+1}, \ldots, \lambda_n$ exactly $p$ have positive real part and $q$
have negative real part.

Proof. We can assume $p + q > 0$, since the assertion is trivial for $D = 0$. The difference $D$ is a solution of the quadratic equation

$$DB_1 + B_1^*D + DCD = 0.$$ 

Moreover $(B_1, C)$ is also controllable. If $\xi \in N$ then, on postmulti-
tiplying (22) by $\xi$, we obtain $DB_1\xi = 0$. Thus $N$ is invariant under

$B_1$. Since $B_2 = B_1 + CD$ it is obvious that $B_1 = B_2$ on $N$.

By replacing $D, B_1, C$ by $T^*DT, T^{-1}B_1T, T^{-1}CT^{-1}$, which does not alter the hypotheses, we can assume that

$$D = \begin{pmatrix} 0 & 0 \\ 0 & D_1 \end{pmatrix},$$

where $D_1$ is non-singular. The nullspace $N$ now consists of all vectors

whose last $p + q$ coordinates vanish. Let the corresponding decompositions of $B_1$ and $C$ be

$$B_1 = \begin{pmatrix} B_{00} & B_{01} \\ 0 & B_{11} \end{pmatrix}, \quad C = \begin{pmatrix} C_{00} & C_{01} \\ C_{10} & C_{11} \end{pmatrix}.$$

Then, putting $G = D_1^{-1}$, we have

$$B_{11}G + GB_{11}^* + C_{11} = 0.$$ 

Moreover
Since $B_{11} + C_{11}D_1 = -GB_{11}^{-1}$, this proves the assertion concerning the
eigenvalues of $B_2$.

We are going to show now that $B_{11}$ has no pure imaginary eigenvalues.
Suppose on the contrary that $B_{11}^{*} \xi_1 = i\mu \xi_1$, where $\mu$ is real and
$\xi_1 \neq 0$. Then from (23) we obtain
$$\xi_1^{*}C_{11}\xi_1 = 0.$$  
Since $C \leq 0$ this implies $C_{11}\xi_1 = 0$, and also $C_{01}\xi_1 = 0$. Thus if we
put
$$\xi = \begin{bmatrix} 0 \\ \xi_1 \end{bmatrix}$$
then
$$\xi^{*}B_{11} = -i\mu \xi^{*}, \quad \xi^{*}C = 0.$$  
Hence
$$\xi^{*}(C_{11} \ldots B_{n-1}C) = 0.$$  
Since $\xi \neq 0$, this contradicts the hypothesis of controllability.

Put
$$B_{11}(\epsilon) = B_{11} + \epsilon G.$$  
Then, by (23), $B_{11}(\epsilon)$ is a solution of the linear equation
$$B_{11}(\epsilon)G + GB_{11}^{*}(\epsilon) + C_{11} - 2\epsilon \xi_1^{*} = 0.$$  
By a generalisation of Ljapunov's Lemma (see Ostrowski and Schneider [10],
Lancaster [6]) it follows that for $\epsilon > 0$, $B_{11}(\epsilon)$ has no pure imaginary
eigenvalues and further that it has $p$ eigenvalues with positive real
part, $q$ eigenvalues with negative real part. Since $B_{11}$ has no pure
imaginary eigenvalues, the same holds by continuity for it.

COROLLARIES. (i) If \( \tilde{W} \) is a solution of (3) such that all eigenvalues of \( B + C\tilde{W} \) have nonpositive real part then \( \tilde{W} = W^+ \). (Take \( W_2 = W^+ \), \( W_1 = \tilde{W} \).)

(ii) If \( \tilde{W} \) is any solution of (3) then the nullspace \( V_0 \) of \( W^+ - W^- \) is invariant under \( B + C\tilde{W} \) and the restriction of \( B + C\tilde{W} \) to \( V_0 \) has only pure imaginary eigenvalues. No other eigenvalues of \( B + C\tilde{W} \) are pure imaginary.

Proof. If \( \xi \in V_0 \) then \( W^+ \xi = \tilde{W}^\ast \xi \), since

\[
0 \leq \xi^* (W^+ - \tilde{W}) \xi \leq \xi^* (W^+ - W^-) \xi = 0.
\]

Since \( V_0 \) is invariant under \( B + CW^+ \) it is equally invariant under \( B + C\tilde{W} \). By Theorem 1 the restriction of \( B + CW^+ \) to \( V_0 \) has only pure imaginary eigenvalues. By Theorem 2, \( B + CW^+ \) has no other pure imaginary eigenvalue. If \( B + C\tilde{W} \) had an additional pure imaginary eigenvalue then, by Theorem 2 again, it would be associated with the nullspace of \( W^+ - \tilde{W} \) and therefore be an eigenvalue of \( B + CW^+ \).

The second corollary has been stated without proof by Willems [15].

It will now be shown that all solutions of (3) can be expressed in terms of the extreme solutions. Put

\[
\Delta = W^+ - W_-, \quad B^+ = B + CW^+, \quad B^- = B + CW^-.
\]

By (4) we have

\[
B^+ \Delta = -\Delta B^-.
\]

The nullspace \( V_0 \) of \( \Delta \) is invariant under \( B^+ \) and, by what we have just proved, there is a unique subspace \( V^+ \) invariant under \( B^+ \) such that the whole space \( V = \mathbb{R}^n \) satisfies

\[
V = V_0 \oplus V^+.
\]
All eigenvalues of the restriction of $B_+$ to $V_0$ are pure imaginary and all eigenvalues of the restriction of $B_+$ to $V_+$ have negative real part.

**Lemma 3.** Let $V_1$ be a subspace invariant under $B_+$. If $V_0 \cap V_1 = 0$ then $V_1 \subseteq V_+$.

Proof. Suppose $\xi \neq 0 \in V_1$. Then $\xi = \xi_0 + \xi_+$, where $\xi_0 \in V_0$ and $\xi_+ \in V_+$. Since $V_0 \cap V_1 = 0$ we must have $\xi_+ \neq 0$. Let $p(s)$ be the polynomial of least positive degree such that $p(B_+)\xi_+ = 0$. Then $p(s)$ has only roots with negative real part and

$$p(B_+)\xi = p(B_+)\xi_0.$$  

Since $V_0$ is invariant under $B_+$ and $V_0 \cap V_1 = 0$ it follows that $p(B_+)\xi_0 = 0$. Therefore $\xi_0 = 0$, and $V_1 \subseteq V_+$.

**Theorem 3.** Let $V_1$ be an arbitrary subspace of $V_+$ invariant under $B_+$. If $V_2$ is the subspace of all vectors $\xi$ such that $A\xi$ is orthogonal to $V_1$ then

$$V = V_1 \oplus V_2.$$ 

Let $P$ be the corresponding projection of $V$ onto $V_1$. Then

$$W = W_1 P + W_2 (I-P)$$  

is a solution of the quadratic equation (3). Moreover all solutions of (3) are obtained in this way, and the correspondence between $V_1$ and $W$ is one-to-one.

Proof. Evidently $V_0 \subseteq V_2$. If $\xi \in V_1 \cap V_2$ then $\xi^*A\xi = 0$. Since $A \geq 0$ it follows that $A\xi = 0$. Since $V_0 \cap V_1 = 0$ this implies $\xi = 0$. Thus $V_1 \cap V_2 = 0$. We have also

$$\dim A \mathbf{V} = n - \dim V_0,$$

$$\dim V_1 = n - \dim V_1.$$

Downloaded from https://www.cambridge.org/core. IP address: 54.70.40.11, on 12 Jun 2019 at 10:02:50, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0004972700041071
and hence
\[ \dim(\Delta U \cap U_1^\perp) \geq n - \dim U_0 - \dim U_1. \]

Therefore
\[ \dim U_2 = \dim U_0 + \dim(\Delta U \cap U_1^\perp) \geq n - \dim U_1. \]

It follows that
\[ U_1^\perp + U_2 = U. \]

If \( \xi_1 \in U_1 \) and \( \xi_2 \in U_2 \) then, by (24),
\[ \xi_1 A B \xi_2 = -\xi_1 A B \xi_2 = 0, \]

since \( U_1 \) is invariant under \( B_+ \). Thus \( U_2 \) is invariant under \( B_- \).

Hence the projection \( P \) satisfies
\[ \text{(27)} \quad P B_+ P = B_+ P \]

and
\[ \text{(28)} \quad (I-P)B_-(I-P) = B_-(I-P); \]

that is,
\[ \text{(29)} \quad PB_+ P = PB_+. \]

Since \( \Delta U_2 \) is orthogonal to \( U_1 \), we have also
\[ P^* \Delta (I-P) = 0; \]

that is,
\[ P^* \Delta = P^* \Delta P. \]

Since the right side is symmetric this gives
\[ \text{(29)} \quad P^* \Delta = \Delta P. \]

By (27) and (28),
\[ (I-P)B_+ = (I-P)B_+(I-P) \]
\[ = (I-P)(B_+ C \Delta)(I-P) \]
\[ = B_-(I-P) + (I-P)C \Delta (I-P). \]

Therefore, by (24),
\[ \Delta(I-P)B_+ + B^*_+\Delta(I-P) = \Delta(I-P)CA(I-P). \]

If we define \( W \) by (26) then
\[ D = W_+ - W = \Delta(I-P). \]

By (29), \( D \) is symmetric, and hence \( W \) is also. Moreover \( W \) is a solution of the quadratic equation (3) if \( D \) is a solution of the quadratic equation
\[ DB_+ + B^*_+D - DCD = 0. \]

But we have just shown that this is the case.

Furthermore
\[ W_\xi = W_+\xi \quad \text{for} \quad \xi \in V_1, \quad W_\xi = W_-\xi \quad \text{for} \quad \xi \in V_2. \]

Therefore \( V_1 \) and \( V_2 \) are invariant under \( B + CW \) and all eigenvalues of the restriction of \( B + CW \) to \( V_1, V_2 \) have respectively negative, non-negative real parts. This proves that \( V_1 \) is uniquely determined by \( W \).

Conversely, let \( W \) be any solution of (3). Then we have a unique direct decomposition
\[ V = V_1 \oplus V_2, \]
where \( V_1 \) and \( V_2 \) are invariant under \( B + CW \) and all eigenvalues of the restriction of \( B + CW \) to \( V_1, V_2 \) have respectively negative, non-negative real parts. It follows from Theorem 2 that
\[ W_\xi = W_+\xi \quad \text{for} \quad \xi \in V_1, \quad W_\xi = W_-\xi \quad \text{for} \quad \xi \in V_2. \]

Thus if \( P \) is the corresponding projection of \( V \) onto \( V_1 \) then \( W \) satisfies (26). Moreover \( V_1 \) is invariant under \( B_+ \) and \( V_0 \cap V_1 = 0. \)

Therefore, by Lemma 3, \( V_1 \subseteq V_+ \). Since \( W_+ - W = \Delta(I-P) \) is symmetric, (29) holds. Therefore, since \( P \) is a projection,
\[ P^*\Delta = P^*\Delta P; \]
that is,
\[ P^*\Delta(I-P) = 0. \]
Thus $\Delta V_2$ is orthogonal to $V_1$. Since $V = V_1 \oplus V_2$ it follows from the first part of the proof that $V_2$ is the subspace of all $\xi$ such that $\Delta \xi$ is orthogonal to $V_1$. This completes the proof.

In particular, if $\Delta > 0$ then $V_0 = 0$, $V_1$ is an arbitrary invariant subspace of $B^+$, and $V_2 = \Delta^{-1} V_1$. Theorem 3 was proved for this case by Willems [15].

In the statement and proof of the following result we use the notation of Theorem 3.

**Theorem 4.** Let $\tilde{W}, \tilde{W}$ be solutions of (3) corresponding to the invariant subspaces $V_1, \tilde{V}_1$ of $B^+$. Then $W \geq \tilde{W}$ if and only if $\tilde{V}_1 \subseteq V_1$.

**Proof.** Suppose first that $\tilde{V}_1 \subseteq V_1$. Then, by Theorem 3, $V_2 \subseteq \tilde{V}_2$. On $\tilde{V}_1$ we have $W \xi = W \xi = \tilde{W} \xi$. On $V_2$ we have $W \xi = W \xi = \tilde{W} \xi$. Therefore $\tilde{V}_1 \oplus V_2$ is contained in the nullspace $N$ of $W - \tilde{W}$ and is invariant under $B + CW$. The remaining eigenvalues of $B + CW$ have negative real part, whereas the remaining eigenvalues of $B + CW$ have non-negative real part. Hence $N = \tilde{V}_1 \oplus V_2$. It follows from Theorem 2 that $W \geq \tilde{W}$.

Conversely, suppose $W \geq \tilde{W}$. The nullspace $N$ of their difference is invariant under $B + CW$, and hence has a unique direct decomposition $N = N_+ \oplus N_-$, where $N_+, N_-$ are invariant under $B + CW$ and all eigenvalues of the restriction of $B + CW$ to $N_+, N_-$ have respectively negative, non-negative real part. By the proof of Lemma 3 we have $N_+ \subseteq V_1$, $N_+ \subseteq \tilde{V}_1$. By Theorem 2 the eigenvalues of $B + CW$ not belonging to $N$ have positive real part. Therefore $\dim N_+ = \dim \tilde{V}_1$. Thus $N_+ = \tilde{V}_1$ and $\tilde{V}_1 \subseteq V_1$.

It follows at once that the set of all solutions of (3) is a complete lattice. This is rather remarkable, since the set of all symmetric $n \times n$
matrices is not a lattice if \( n > 1 \).

3. Existence of solutions

In this section we drop the hypotheses and convention imposed at the beginning of Section 2. If \( \tilde{W} \) is a solution of (3), not assumed symmetric, such that \( B + \tilde{C}W \) is stable then the characteristic polynomial

\[
\psi(s) = \det(sI-B-C\tilde{W})
\]

is relatively prime to \( \psi(-s) \). Since, by (9), \( \psi(s) \) divides \( \varphi(s) = \det(sI-M) \), and \( \varphi(s) \) is even, it follows that

\[
\varphi(s) = (-1)^n\psi(s)\psi(-s) .
\]

On the other hand it is clear that \( \varphi(s) \) admits a factorisation (30), where the (real) polynomial \( \psi(s) \) is monic and relatively prime to \( \psi(-s) \), if and only if \( M \) has no pure imaginary eigenvalue. Moreover it is possible to choose \( \psi(s) \) so that all its roots lie in the left half-plane.

**Theorem 5.** Suppose \( \varphi(s) \) has a factorisation (30), where \( \psi(s) \) is monic and relatively prime to \( \psi(-s) \). Then there exists a non-singular matrix \( X \) such that

\[
X^*JX = J , \tag{31}
\]

\[
MX = \begin{bmatrix} T & 0 \\ 0 & -T^* \end{bmatrix} , \tag{32}
\]

\[
\det(sI-T) = \psi(s) . \tag{33}
\]

Proof. There exist polynomials \( \chi_1(s) , \chi_2(s) \) such that

\[
1 = \chi_1(s)\psi(s) + \chi_2(s)\psi(-s) . \tag{34}
\]

Then for any vector \( x \in \mathbb{R}^{2n} \),

\[
x = \chi_1(M)\psi(M)x + \chi_2(M)\psi(-M)x . \tag{35}
\]

Let \( V_1 \) denote the subspace of all vectors \( x \) such that \( \psi(M)x = 0 \), and let \( V_2 \) denote the subspace of all vectors \( x \) such that \( \psi(-M)x = 0 \).

\[3\text{ A matrix } X \text{ satisfying (31) is said to be symplectic.}\]
From (35) we obtain at once

\[ V_1 \cap V_2 = 0 , \]

\[ V_1 + V_2 = R^{2n} . \]

Moreover \( V_1 \) and \( V_2 \) are invariant under \( M \).

Let \( x_1, \ldots, x_m \) be a basis for \( V_1 \) and \( x_{m+1}, \ldots, x_{2n} \) a basis for \( V_2 \), and let \( \tilde{X} \) be the non-singular matrix whose \( k \)-th column is \( x_k \) \((k = 1, \ldots, 2n) \). Then

\[ \tilde{X}M = \begin{bmatrix} T & 0 \\ 0 & S \end{bmatrix} , \]

where \( T \) is an \( m \times m \) matrix, \( S \) is a \((2n-m) \times (2n-m)\) matrix, and \( \psi(T) = 0 \), \( \psi(-S) = 0 \). Thus the minimal polynomial of \( T \) divides \( \psi(s) \). Since every prime factor of the characteristic polynomial of a matrix divides its minimal polynomial, it follows that \( \det(sI - T) \) is relatively prime to \( \psi(-s) \). On the other hand, \( \det(sI - T) \) divides \( \phi(s) = (-1)^n \psi(s) \psi(-s) \). Therefore \( \det(sI - T) \) divides \( \psi(s) \). Thus \( m > 0 \) implies \( m \leq n \). Similarly \( m < 2n \) implies \( m \geq n \). Hence \( m = n \) and (33) holds.

It follows from (34) that \( \psi(-T) \) and \( \psi(S) \) are non-singular. Also, by (7),

\[ \tilde{X}J\tilde{X} \begin{bmatrix} T & 0 \\ 0 & S \end{bmatrix} = \tilde{X}JM\tilde{X} \]

\[ = \tilde{X}M^{*}J\tilde{X} \]

\[ = \begin{bmatrix} -T^{*} & 0 \\ 0 & -S^{*} \end{bmatrix}\tilde{X}J\tilde{X} . \]

Since \( \psi(T) = 0 \), \( \psi(-S) = 0 \) it follows that

\[ \tilde{X}J\tilde{X} \begin{bmatrix} 0 & 0 \\ 0 & \psi(S) \end{bmatrix} = \begin{bmatrix} \psi(-T)^{*} & 0 \\ 0 & 0 \end{bmatrix}\tilde{X}J\tilde{X} . \]

Write, in partitioned form,
Matrix quadratic equations

\[ X^*JX = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} , \]

where \( U_{12} = - U_{21} \). Then (37) implies that \( U_{11} = 0 \), \( U_{22} = 0 \). Since \( X^*JX \) is non-singular it follows that \( U_{12} \) is non-singular. From (36) we now obtain

\[ U_{12}S = - T^A U_{12} . \]

Put

\[ X = \begin{bmatrix} I & 0 \\ 0 & -U^{-1} \end{bmatrix} . \]

Then (31) and (32) hold. This completes the proof.

Write in partitioned form

\[ X = \begin{bmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{bmatrix} . \]

Then (31) is equivalent to

\[ Z^*Y_1 = Y_1^*Z_1 , \]

(38)

\[ Z^*Y_2 = Y_2^*Z_2 , \]

\[ Z_1^* - Y_2^*Z_1 = I , \]

and (32) implies

\[ BY_1 + CZ_1 = Y_1^* , \]

(39)

\[ -AY_1 - B^*Z_1 = Z_1^* . \]

Suppose \( Y_1 \) is non-singular, and put \( W_1 = Z_1^* Y_1^{-1} \). Then, by (38), \( W_1 \) is symmetric. By (39),

\[ B + CW_1 = Y_1^* Y_1^{-1} . \]
$A + \hat{W}B = -Y^{-1}_1 T^*Z^*_1$.

Hence $\det(sI-B-C\hat{W}) = \psi(s)$ and

$$
Q[\hat{W}] = A + \hat{W}B + (B^*+\hat{W}^*C)\hat{W}
= -Y^{-1}_1 T^*Z^*_1 + \left(Y^{-1}_1 T^*Y^*_1\right)\hat{W}
= 0 .
$$

Let $\hat{W}$ be any solution of (3) such that $\psi(s) = \det(sI-B-C\hat{W})$. Then

$$
M \begin{bmatrix} I \\ \hat{W} \end{bmatrix} = M \begin{bmatrix} I \\ \hat{W} \end{bmatrix} (B+CW)
$$

and hence

$$
\psi(M) \begin{bmatrix} I \\ \hat{W} \end{bmatrix} = \psi(M) \begin{bmatrix} I \\ \hat{W} \end{bmatrix} (B+CW) = 0 .
$$

Consider in general the linear equation

(40)

$$
\psi(M) \begin{bmatrix} I \\ \hat{W} \end{bmatrix} = 0 .
$$

Writing $\psi(M) = \psi(M)X.K^{-1}$, where

$$
\psi(M)X = X \begin{bmatrix} 0 & 0 \\ 0 & \psi(-T)^* \end{bmatrix}
$$

and

$$
X^{-1} = J^{-1}X^*J = \begin{bmatrix} Z^*_1 & -Y^*_2 \\ -Z^*_2 & Y^*_1 \end{bmatrix},
$$

and recalling that $\psi(-T)$ is non-singular, we see that (40) is equivalent to

$$
\hat{W}^*Y_1 = Z_1 .
$$

Since $\psi$ is non-singular there is no vector $\xi \neq 0$ such that $Y_1\xi = Z_1\xi = 0$. It follows that (40) is soluble only if $Y_1$ is non-singular, and it then has the unique solution $\hat{W} = \hat{W}_1$. 


Thus the equation (3) has a solution $W_1$ such that
$$\psi(s) = \det(sI-B-CW_1)$$
if and only if $Y_1$ is non-singular. When such a solution exists, it is unique and symmetric. Similarly the equation (3) has a solution $W_2$ such that $(-1)^n\psi(-s) = \det(sI-B-CW_2)$ if and only if $Y_2$ is non-singular. Moreover $W_2$ is then the unique solution of the linear equation
$$(-W I)\psi(M) = 0,$$
with the same coefficient matrix as (40). If $W_1$ and $W_2$ both exist then $W_2 - W_1$ is non-singular, since by (38),
$$W_2 - W_1 = y_2^{-1}y_1^{-1}.$$
Proof. Let

\[ X = \begin{pmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{pmatrix} \]

be the matrix whose existence was established in Theorem 5. We have already seen that if \( Y_1 \) is non-singular then (ii) holds.

Suppose \( Y_1 \) is singular, and let \( \xi \in \mathbb{R}^n \) be any vector such that \( Y_1 \xi = 0 \). From (39) and (38) we obtain

\[ \xi^*Z_1BY_1\xi + \xi^*Z_1^*C_1\xi = \xi^*Z_1^*Y_1T\xi \]

Since \( Y_1 \xi = 0 \) it follows that \( \xi^*Z_1^*C_1\xi = 0 \). Therefore, since \( R > 0 \),

(41) \[ E^*Z_1\xi = 0. \]

Putting \( \eta = Z_1\xi \), we now obtain from (39),

(42) \[ Y_1T\xi = 0, \]

\[ -B^*\eta = Z_1T\xi. \]

This shows that we can replace \( \xi \) by \( T\xi \) in the preceding argument. Thus for any non-negative integer \( k \),

(43) \[ E^*Z_1T^k\xi = 0, \]

\[ Y_1T^k\xi = 0. \]

For any \( m \times n \) matrix \( F \),

\[ \eta^*(B+EF) = -\xi^*T^*Z_1^* + \xi^*Z_1^*EF \]

\[ = -\xi^*T^*Z_1^*. \]

By induction on \( k \), using (43) and (39), it is easily shown that

\[ \eta^*(B+EF)^k = (-1)^k\xi^*(T^*)^kZ_1^*. \]

Hence for any polynomial \( p(s) \),
Matrix quadratic equations

\[ \eta^*p(-B-\mathcal{E}) = \xi^*p(T^*)Z_1^* \, . \]

If \( Z_1 p(T) \xi = 0 \) then \( p(T) \xi = 0 \), since \( Y_1 p(T) \xi = 0 \) and \( X \) is non-singular.

Fix a particular \( \xi \neq 0 \) in the nullspace of \( Y_1 \) and let \( \omega(s) \) be the monic polynomial of least positive degree such that \( \omega(T) \xi = 0 \). Then \( \omega(s) \) divides the minimal polynomial of \( T \), and hence also \( \psi(s) \). Furthermore \( \omega(s) \) is the monic polynomial of least degree such that

\[ \eta^*\omega(-B-\mathcal{E}) = 0 \, . \]

Thus \( \omega(-s) \) divides the minimal polynomial, and also the characteristic polynomial, of \( B + \mathcal{E} \). This proves (i). Since

\[ B + \mathcal{E} = B + E \left[ -R^{-1}E^* \mathcal{W}_1 \right] \, , \]

and \( \psi(s) \) is relatively prime to \( \psi(-s) \), it is clear that (i) and (ii) cannot hold simultaneously.

By replacing \( \psi(s) \) by \( (-1)^n \psi(-s) \) and \( Y_1, Z_1 \) by \( Y_2, Z_2 \) we obtain (i)' and (ii)'. Suppose now that \( (B, \mathcal{E}) \) is controllable. Then by a well-known property of controllability (see, for example, Heymann [2]), neither (i) nor (i)' holds. Therefore both (ii) and (ii)' hold.

Conversely, suppose that \( (B, \mathcal{E}) \) is not controllable. Then there exists a non-singular \( n \times n \) matrix \( P \) such that

\[ PBP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}, \quad PE = \begin{pmatrix} E_1 \\ 0 \end{pmatrix}, \]

where \( \dim B_{22} \geq 1 \). Write

\[ P = [P_1 \, P_2]P \, . \]

Then

\[ P(B + \mathcal{E})P^{-1} = \begin{pmatrix} B_{11} & E_1 P_{11} & B_{12} & E_1 P_{12} \\ 0 & B_{22} \end{pmatrix}, \]

and hence
If both (ii) and (ii)' hold then, taking \( F = -R^{-1}E^{*}W_i \) \((i = 1, 2)\), we see that \( \det(sI-B_{22}) \) divides both \( \psi(s) \) and \( \psi(-s) \). Since they are relatively prime and \( \det(sI-B_{22}) \) is of positive degree this is a contradiction.

**COROLLARY.** Suppose \( C \leq 0 \). Then the following statements are equivalent:

(i) \((B, C)\) is controllable and \( M \) has no pure imaginary eigenvalues;

(ii) the equation \((3)\) has a solution \( W_1 \) such that \( B + CW_1 \) is stable and a solution \( W_2 \) such that \( -(B + CW_2) \) is stable.

This Corollary has been established by Molinari [8], [9] by a rather long argument involving spectral factorisation of matrices of rational functions. The key idea in the proof of Theorem 6, namely, the derivation of (41) and (42), is due to Kucera [5]. However Kucera treats only a special case and uses the Jordan canonical form.

We next use Theorem 5 to discuss the applicability of an algorithm introduced by Roberts [13].

**LEMMA 4.** Let \( A \) be a linear transformation of the vector space \( V \) which has no pure imaginary eigenvalues, and let \( V = V_+ \oplus V_- \) be the direct decomposition of \( V \) into subspaces invariant under \( A \) such that the eigenvalues of the restriction of \( A \) to \( V_-, V_+ \) have respectively negative, positive real parts. Set

\[
A_0 = A, \quad A_{v+1} = \left(A_v + A_v^{-1}\right)/2 \quad (v \geq 0).
\]

Then the sequence \( \{A_v\} \) is defined for all \( v \) and \( A_v \to A_\omega \) as \( v \to \infty \), where \( A_\omega \) is the linear transformation defined by

\[
A_\omega \xi = -\xi \text{ if } \xi \in V_-, \quad A_\omega \xi = \xi \text{ if } \xi \in V_+.
\]
Proof. Since the effect of changing $A$ into $-A$ is obvious we may suppose that $A$ is stable. If $A$ has the eigenvalues $\lambda_k$ $(k = 1, \ldots, n)$ then

$$\tilde{A} = (A + A^{-1})/2$$

has the eigenvalues

$$\tilde{\lambda}_k = \left(\lambda_k + \lambda_k^{-1}\right)/2 \quad (k = 1, \ldots, n).$$

Hence $\tilde{A}$ is also stable. It follows that $A - I$ and $\tilde{A} - I$ are nonsingular and

$$(\tilde{A} + I)(\tilde{A} - I)^{-1} = \left[(A + I)(A - I)^{-1}\right]^2.$$

Therefore the sequence $\{A_\nu\}$ is defined for all $\nu$ and

$$\left(A_\nu + I\right)\left(A_\nu - I\right)^{-1} = \left[(A + I)(A - I)^{-1}\right]^{2^\nu}.$$

Since all eigenvalues of $(A + I)(A - I)^{-1}$ have absolute value less than 1 this shows that $\left(A_\nu + I\right)\left(A_\nu - I\right)^{-1} \to 0$, and hence $A_\nu \to -I$.

It is easily seen that if $A$ has at least one pure imaginary eigenvalue then the sequence $\{A_\nu\}$ is either not defined for all $\nu$ or $\lim_{\nu \to \infty} A_\nu$ does not exist.

Suppose now that the matrix $M$ in (6) has no pure imaginary eigenvalues. Then, by Theorem 5, there exists a symplectic matrix $X$ such that

$$M = X\begin{bmatrix} T & 0 \\ 0 & -T^* \end{bmatrix}X^{-1},$$

where $T$ is stable. Hence

$$M_0 = X\begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix}X^{-1}.$$

Writing
we get
\[
\hat{A} = Z_1^*Z_2 + Z_2^*Z_1 = 2Z_2^*Z_1 = 2Z_1^*Z_2,
\]
\[
\hat{B} = -Y_1Z_2^* - Y_2Z_1^* = -2Y_1Z_2^* = I = -2Y_1Z_2 + I,
\]
\[
\hat{C} = Y_1^*Y_2 + Y_2^*Y_1 = 2Y_1^*Y_2 = 2Y_2^*Y_1.
\]
If \(Y_1\) and \(Y_2\) are non-singular then \(\hat{C}\) is non-singular and the linear equations
\[
(46) \quad \hat{C}W = -I - \hat{B}, \quad \hat{C}W = I - \hat{B}
\]
have the unique solutions \(W_1 = Z_1^*Y_1^{-1}\), \(W_2 = Z_2^*Y_2^{-1}\). Thus whenever (3) has both stabilizing and antistabilizing solutions they can be found by Roberts' algorithm. This result is independent of whether \(C \leq 0\), and consequently should be useful in the applications to differential games.

References


Department of Mathematics,
Institute of Advanced Studies,
Australian National University,
Canberra, ACT.