CONDITIONS FOR A MATRIX TO COMMUTE WITH ITS INTEGRAL

IRVING J. EPSTEIN

1. Introduction. Let $U(t)$ be an $n \times n$ matrix whose elements are continuous functions of a parameter $t$. We shall find necessary and sufficient conditions for the relation

$$U(t) \int_0^t U(s)ds = \left( \int_0^t U(s)ds \right) U(t)$$

(1.1)

to hold in an interval $0 \leq t \leq t_0$, where $t_0$ is so small that throughout the interval $[0, t_0]$ the Jordan canonical form of $U(t)$ has the same form. That is, its off-diagonal elements do not change in the interval.

Matrices $U(t)$ satisfying (1.1) are of interest for various reasons; see, for instance, [1, p. 278]. We may mention two occasions where (1.1) occurs. Firstly, consider a system of $n$ homogeneous linear differential equations of the first order for $n$ unknown functions with $U(t)$ as the matrix of coefficients. If we consider the unknown functions as components of a vector, and if we form a matrix $Y$, the $n$ columns of which are $n$ linearly independent solutions of our system, then we have for $Y = Y(t)$:

$$\dot{Y} = UY, \quad Y(0) = I,$$

(1.2)

where a dot denotes the derivative with respect to $t$ and where $I$ denotes a nonsingular matrix which we may choose to be the unit matrix. If (1.1) holds, then (1.2) can be solved in terms of quadratures. In fact, we have

$$Y = \exp \int_0^t U(s)ds.$$

(1.3)

Secondly, consider the following problem in the theory of systems with periodic coefficients. Let $W(t)$ be an $n \times n$ matrix depending continuously on $t$ such that, for a constant $\omega$,

$$W(t + \omega) = W(t),$$

(1.4)

and also

$$W(-t) = -W(t).$$

(1.5)

Then it has been shown in special cases by Demidovic [2] and, more
generally, by the author [3] that a matrix $Z_0(t)$ satisfying
\begin{equation}
\dot{Z}_0 = WZ_0, \quad Z_0(0) = I,
\end{equation}
is periodic with period $\omega$, i.e., $Z_0(t+\omega) = Z_0(t)$. The matrices $W(t)$ form a linear space under addition.

We ask whether we can extend this linear space such that a system of the type (1.6) still will have periodic solutions. A partial answer to this question is given by the following remark: Let $W$ be a fixed matrix satisfying (1.4) and (1.5). Let $E(t)$ be such that
\begin{equation}
E(t + \omega) = E(t), \quad E(-t) = E(t).
\end{equation}
Then the system
\begin{equation}
\dot{Z} = (W + \epsilon E)Z, \quad Z(0) = I
\end{equation}
will have solutions with the property
\begin{equation}
Z(t + \omega) = Z(t),
\end{equation}
for all values of $\epsilon$ if the matrix $A(t)$ defined by
\begin{equation}
A(t) = Z_0^{-1}(t)E(t)Z_0(t)
\end{equation}
commutes with its integral. The proof is based on the standard procedure of expanding $Z(t)$ in a power series in $\epsilon$.

2. Matrices commuting with their derivatives. Instead of $U(t)$ we introduce
\begin{equation}
V(t) = \int_0^t U(s) \, ds,
\end{equation}
and assume that
\begin{equation}
\dot{V}V - V\dot{V} = 0.
\end{equation}
Consider an interval $(t_1, t_2)$ such that, for $t_1 \leq t \leq t_2$, there exists a differentiable nonsingular matrix $P(t)$ such that
\begin{equation}
V(t) = P^{-1}(t)J(t)P(t),
\end{equation}
where $J(t)$ is in Jordan canonical form. This means that
\begin{equation}
J = \begin{bmatrix}
C_1 & 0 & \cdots & 0 \\
0 & C_2 & \cdots & \\
& & \ddots & \\
0 & \cdots & \cdots & C_r
\end{bmatrix},
\end{equation}
where the submatrices $C_p(t)$, $p = 1, \ldots, r$, are $n_p \times n_p$ matrices of the form.
(2.4) \[ C_p = \alpha_p(t)I_p + \delta_p E_p. \]

Here \( \alpha_p(t) \) is a differentiable function of \( t \), \( I_p \) is the \( np \times np \) unit matrix, \( \delta_p \) is 0 or 1, and \( E_p \) is the \( np \times np \) matrix with elements \[ e_{\mu,\nu}, \quad \nu, \mu = 1, \cdots, np, \]

and
\[
(2.5) \quad e_{\mu,\nu+1} = 1, \quad e_{\mu,\mu} = 0 \quad \text{for} \quad \mu - \nu \neq 1.
\]

We shall assume that the interval \( (t_1, t_2) \) is such that no difference \( \alpha_p - \alpha_\sigma \) vanishes in a subinterval unless it vanishes identically.

We may assume that, if \( \alpha_p - \alpha_\sigma \) vanishes identically, for \( \rho \neq \sigma \), either \( \delta_\rho \neq 0 \) or \( \delta_\sigma \neq 0 \). Otherwise, we could contract \( C_p \) and \( C_\sigma \) into a single diagonal matrix.

Now we have:

**Theorem 1.** The general \( n \times n \) matrix \( V(t) \) satisfying (2.1) and having a Jordan canonical form determined by (2.3), (2.4) with constant \( np, \delta_p \) for \( t_1 \leq t \leq t_2 \) is obtained by finding all \( n \times n \) matrices \( X \) satisfying
\[
(2.6) \quad J(XJ - JX) - (XJ - JX)J = 0,
\]

determining the nonsingular solutions \( P(t) \) of the matrix differential equation
\[ P = XP, \]
and forming
\[ V = P^{-1}JP. \]

The matrices \( X \) form a linear space (under addition) which depends only on the \( np, \delta_p \), and the set of pairs of subscripts \( (\rho, \sigma) \) for which \( \alpha_\rho - \alpha_\sigma \) vanishes identically.

**Proof.** We observe that, trivially,
\[
(2.7) \quad JJ = \dot{J}J.
\]

By differentiating (2.2), we find
\[
(2.8) \quad \dot{V} = P^{-1}JP + P^{-1}JP + P^{-1}JP.
\]

Because of \( P^{-1}P = I \) we have
\[ \dot{P}^{-1}P + P^{-1}P = 0, \quad \dot{P}^{-1} = -P^{-1}PP^{-1}, \]
and therefore from (2.2), (2.8), with \( X = \dot{P}P^{-1} \):
\[ \dot{VV} - VV = P^{-1}\{-XJ + J + JX\}JP - P^{-1}\{-XJ + J + JX\}P = 0. \]
If we multiply this last equation by \( P \) on the left and \( P^{-1} \) on the right and then make use of (2.7) we get (2.6). We note that the solutions \( X \) of (2.6) form a linear space. In the next section, we shall determine a basis for the linear space of the matrices \( X \) and, incidentally, shall also prove that this space does not depend on the functions \( \alpha_p(t) \) but merely on the discrete parameters mentioned in Theorem 1.

**Corollary.** A system of linear differential equations which, in matrix form, can be written as

\[
\dot{Y} = UY
\]

where the coefficient matrix \( U = \dot{V} \) has the property \( UV = VU \), can always be transformed into a system

\[
\dot{Z} = (X + J + JX - XJ)Z,
\]

where \( X, J \) are defined as in Theorem 1. The transformation to be used is, of course, \( Z = PY \), where \( P \) is defined as in Theorem 1.

3. The space of matrices \( X \). The solutions \( X \) of (2.6) may be written as matrices which are composed of submatrices \( X_{p,\sigma}, p, \sigma = 1, \ldots, \tau \), where \( X_{p,\sigma} \) is a matrix with \( n_p \) rows and \( n_\sigma \) columns and

\[
X = (X_{p,\sigma})
\]

with the natural arrangement of the submatrices. From (2.6) we find the equations

\[
C_p^2 X_{p,\sigma} + X_{p,\sigma} C_\sigma^2 - 2C_p X_{p,\sigma} C_\sigma = 0,
\]

where \( C_p \) is given by equation (2.4).

If we let \( x_{k,l} \) denote the element in the \( k \)th row and \( l \)th column of \( X_{p,\sigma} \) then (3.2) gives us the scalar equations

\[
(a_p - \alpha_\sigma)^2 x_{k,1} + 2\delta_\sigma (a_p - \alpha_\sigma) x_{k,1+1} + 2\delta_\sigma (a_\sigma - \alpha_p) x_{k,1+1} + 2\delta_\sigma x_{k,1-1} - 2\delta_\sigma x_{k,1+1,1-1} = 0
\]

where

\[
k = 1, 2, \ldots, n_p, \quad l = 1, 2, \ldots, n_\sigma,
\]

and where we define \( x_{p,q} = 0 \) if \( p > n_p \) or \( q < 1 \). Equations (3.3) have to be analyzed for various cases. We may summarize the results as follows:

**Theorem 2.** The matrix \( X_{p,\sigma} \) has one of the following structures:

CASE 1. \( \alpha_p - \alpha_\sigma \) does not vanish identically (and, therefore, not in any subinterval of \( (t_1, t_2) \)). Then \( X_{p,\sigma} \) is identically zero.
Case 2. \( \alpha_p - \alpha_q = 0, \delta_p = 0, \delta_q = 1 \). Then the last two columns of \( X_{p,q} \) are arbitrary, but all other elements of \( X_{p,q} \) vanish identically.

Case 3. \( \alpha_p - \alpha_q = 0, \delta_p = 1, \delta_q = 0 \). Then the first two rows of \( X_{p,q} \) are arbitrary but all other elements of \( X_{p,q} \) vanish identically.

Case 4. \( \alpha_p - \alpha_q = 0, \delta_p = \delta_q = 0 \). Then we may assume \( \rho = \sigma \) (see remarks before Theorem 1), and \( X_{p,q} \) is arbitrary.

Case 5. \( \alpha_p - \alpha_q = 0, \delta_p = \delta_q = 1 \). Denoting the elements of \( X_{p,q} \) by \( x_{l,k} \), where \( l = 1, \ldots, n_p \) and \( k = 1, \ldots, n_q \), and if \( n_p > n_q \), then the first two rows of \( X_{p,q} \) are arbitrary and \( X_{p,q} \) has the appearance indicated below:

\[
\begin{array}{ccccccc}
   x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & \cdots \\
   x_{21} & x_{22} & x_{23} & x_{24} & x_{25} & \cdots \\
   0 & 2x_{21} & 2x_{22} - x_{11} & 2x_{23} - x_{12} & 2x_{24} - x_{13} & \cdots \\
   0 & 0 & 3x_{21} & 3x_{22} - 2x_{11} & 3x_{23} - 2x_{12} & \cdots \\
   0 & 0 & 0 & 4x_{21} & 4x_{22} - 3x_{11} & \cdots \\
   0 & 0 & 0 & 0 & 5x_{21} & \cdots \\
   \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\end{array}
\]

If \( n_p < n_q \), the role of rows and columns has to be exchanged, and if \( n_p = n_q \), the \( X_{p,q} \) is triangular, but the same shape as above, except that \( x_{21} = 0 \).

Only Case 5 requires a more detailed analysis. However, once the explicit form of \( X \) stated above is known, it can be verified with a moderate amount of calculations which will be omitted here.

Acknowledgment. I acknowledge with pleasure the advice given me by Professor W. Magnus of New York University in the preparation of this paper.

References

Evans Signal Laboratory, Ft. Monmouth