The Solution of the Matrix Equations $AXB - CXD = E$ AND $(YA - DZ, YC - BZ) = (E, F)$

King-wah Eric Chu
Mathematics Department
University of Reading
P.O. Box 220
Reading, England RG6 2AX

Submitted by Richard A. Brualdi

ABSTRACT

The conditions for the existence of a unique solution of the matrix equation $AXB - CXD = E$ are proved to be that (i) the pencils $A - \lambda C$ and $D - \lambda B$ are regular, and (ii) the spectra of the pencils have an empty intersection. A numerical algorithm for solving the equation is proposed. The possibility of a least-squares-type solution is briefly discussed. The set of equations $(YA - DZ, YC - BZ) = (E, F)$ is proved to be equivalent to the aforementioned equation, and its solution is also investigated. A numerical algorithm is proposed.

1. INTRODUCTION

We consider the matrix equation for $X \in \mathbb{R}^{m \times n}$,

$$AXB - CXD = E, \tag{1}$$

where, $A, C \in \mathbb{R}^{m \times m}$ and $D, B \in \mathbb{R}^{n \times n}$.

Equation (1) is a special case of the general linear equation in X:

$$\sum_{i=1}^{p} A_i X B_i = E. \tag{2}$$

By using the Kronecker tensor product, Equation (2) [and thus (1)] can be

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

written as

$$\left\{\sum_{i=1}^{p}\left(A_i \otimes B_i^T\right)\right\} v(X) = v(E), \tag{3}$$

where $v(X) = (x_1^T, x_2^T, \dots, x_m^T)^T$, with x_i^T the ith row of X. Equation (3) is now a simple set of linear simultaneous equations, with mn equations in mn unknowns. The solvability of Equations (1) and (2) can then be investigated through looking at Equation (3). In [8], this was done for the special case $p = 1$. For $p > 1$, the matrix in Equation (3) has too complicated a structure and no general result for the solvability problem is available. Also, solving Equations (1) and (2) in the form of Equation (3) using the usual Gaussian elimination technique fails to take account of the matrix structure and requires $O(m^3n^3)$ flops. The operation count is obviously unacceptable when $m \cong n$.

In this paper, a set of necessary and sufficient conditions for the existence and uniqueness of the solution of Equation (1) is presented. A numerical algorithm for the solution is proposed, and the problem of ill-conditioning considered. Possibilities of solving Equation (1) in the least-squares sense are discussed briefly in Section 5.

In Section 6, Equation (1) is proved to be equivalent to the set of equations

$$YA - DZ = E,$$

$$YC - BZ = F,$$

and a numerical algorithm is proposed for its solution.

Note that the Equation (1) is a generalization of the Sylvester equation $AX - XD = E$, discussed by Bartels and Stewart [1]; this paper is strongly influenced by their work.

The author came across equations of the type in (1) when analysing perturbation problems of the generalized eigenvalue problem [2]. Another numerical algorithm for the solution of Equation (1) can be found in [4], together with some applications.

2. THE SOLVABILITY OF EQUATION (1)

Consider the generalized eigenvalue problem [2, 5, 7, 9, 11, 12]

$$Ax = \lambda Cx \tag{4}$$

in the more sensible and convenient form

$$\gamma A x = \alpha C x, \tag{5}$$

with some normalization for x, e.g. $\|x\|_2 = 1$. Note that the roles of A and C are now symmetric, and zero and infinite eigenvalues λ will now be treated similarly, as $(\alpha, \gamma) = (0, \gamma)$ or $(\alpha, 0)$. From Equations (4) and (5), one has

$$\lambda = \alpha / \gamma, \tag{6}$$

with $\lambda = \infty$ when $\gamma = 0$.

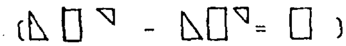
Consider a regular pencil $A - \lambda C$. In general, there exist unitary matrices P_1 and P_2, through the QZ decomposition [7, 11], such that $P_1 A P_2 \triangleq (\alpha_{ij})$ and $P_1 C P_2 \triangleq (\gamma_{ij})$ are both *lower* triangular, with diagonal elements $\{\alpha_{ii}\}$ and $\{\gamma_{ii}\}$ respectively. The generalized eigenvalues will then be $(\alpha, \gamma) = \{\alpha_{ii}, \gamma_{ii}\}$. Note that $\alpha_{ii} = \gamma_{ii} = 0$ is impossible for any i, as it will indicate a singular pencil. Similarly, there exist unitary matrices Q_1 and Q_2 such that $Q_1 D Q_2 \triangleq (\delta_{ij})$ and $Q_1 B Q_2 \triangleq (\beta_{ij})$ are both *upper* triangular, with $D - \lambda B$ a regular pencil (cf. [1]). We defined the spectra $\rho(A, C)$ and $\rho(D, B)$ as the collections of $(\alpha_{ii}, \gamma_{ii})$ and $(\delta_{jj}, \beta_{jj})$ respectively. Use the usual equivalence relation $\equiv$ for quotients, where

$$(\alpha, \gamma) \equiv (\delta, \beta) \quad \text{iff} \quad \alpha\beta - \gamma\delta = 0. \tag{7}$$

From now on, we only consider the equivalence classes in $\rho(A, C)\ \rho(D, B)$.

Equation (1) has now been transformed to

$$\begin{aligned} & P_1 A P_2 \cdot P_2^H X Q_1^H \cdot Q_1 B Q_2 - P_1 C P_2 \cdot P_2^H X Q_1^H \cdot Q_1 D Q_2 = P_1 E Q_2 \\ \Leftrightarrow \quad & \tilde{A}\tilde{X}\tilde{B} - \tilde{C}\tilde{X}\tilde{D} = \tilde{E}\ \left(\triangleq (\varepsilon_{ij}) \right). \end{aligned} \tag{8}$$



Here, $(\cdot)^H$ denotes the Hermitian.

Consider $\tilde{x}_{ij}$, the $(i, j)th$ component of $\tilde{X}$, row-wise. Equation (8) can then be written in the form, with Δ_{ijkl} denoting $\alpha_{ij}\beta_{kl} - \gamma_{ij}\delta_{kl}$,

$$\Delta_{1111} \cdot \tilde{x}_{11} = \varepsilon_{11},$$

$$\Delta_{1122} \cdot \tilde{x}_{12} = \varepsilon_{12} - \Delta_{1112} \cdot \tilde{x}_{11},$$

and for a general (i, j),

$$\Delta_{iijj} \cdot \tilde{x}_{ij} = \varepsilon_{ij} - \sum_{\substack{k=1 \\ (k,l) \neq (i,j)}}^{i} \sum_{l=1}^{j} \Delta_{iklj} \cdot \tilde{x}_{kl}. \tag{9}$$

It is obvious from Equation (9) that Equation (8), and thus (1), can be solved for a unique X, if and only if $\Delta_{iijj} \neq 0 \ \forall i, j$, that is, $\rho(A, C) \cap \rho(D, B) = \varnothing$.

The above argument provides a solution process for Equation (1) and the motivation for the following theorem. The theorem can be proved using a similar argument, but a neater proof is provided.

THEOREM 1. *The matrix equation* (1) *has a unique solution if and only if*

(i) $A - \lambda C$ *and* $D - \lambda B$ *are regular matrix pencils, and*
(ii) $\rho(A, C) \cap \rho(D, B) = \varnothing$.

(Recall the equivalence classes defined by Equation (7).*)*

Proof. Consider the equations

$$(\lambda_1 A - \lambda_2 C)XB - CX(\lambda_1 D - \lambda_2 B) = E \tag{10a}$$

and

$$(\lambda_1 A - \lambda_2 C)XD - AX(\lambda_1 D - \lambda_2 B) = -F, \tag{10b}$$

for some real λ_1 and λ_2 which are not both zero. One of the equations in (10) is equivalent to Equation (1). (Cf. [4].)

If the conditions (i) and (ii) are satisfied, λ_i's can be found so that the matrices involving the λ_i's are nonsingular; thus solving Equation (10a) or (10b) is equivalent to solving an equivalent Sylvester equation, which yields a unique solution.

If either (or both) of the conditions (i) and (ii) are violated, some λ_i's can be found such that the matrices $\lambda_1 A - \lambda_2 C$ and $\lambda_1 D - \lambda_2 B$ are singular. Let $y \neq 0$ and $z \neq 0$ be such that $(\lambda_1 A - \lambda_2 C)y = 0$ and $z^H(\lambda_1 D - \lambda_2 B) = 0$. Then cyz^H, for any nonzero constant c, will be a nontrivial solution of the homogeneous equation related to equation (10a) or (10b). As a result, a solution cannot be unique, if it exists at all. ∎

Note that for the Sylvester equation, with $B = I_m$ and $C = I_n$, the conditions in Theorem 1 reduce to $\rho(A) \cap \rho(D) = \varnothing$. (See [1].)

Note also that the solution process through Equation (9) is equivalent to constructing X from the generalized eigensystems of (A, C) and (D, B). Any violation of conditions (i) and (ii) can then be detected, in theory, by inspecting the spectra $\rho(A, C)$ and $\rho(D, B)$, after the QZ processes have been performed in Equation (8). More discussions on the numerical aspects of ill-conditioning arising from the solution of Equation (1) can be found at the end of Section 3.

Finally, even if the matrices B and C are nonsingular and the equation (1) can be transformed to the Sylvester equation form

$$C^{-1}AX - XDB^{-1} = C^{-1}EB^{-1}, \tag{11}$$

one should not solve (1) in the form of (11). Denoting the equation (1), using the operator T, as

$$T(X) \triangleq AXB - CXD = E, \tag{12}$$

it is easy to see that the conditioning of the solution of (1) can be represented by the condition number $\kappa(T)$, and that of (11) by $\kappa(T)\kappa(B)\kappa(C)$, with $\kappa(T) \triangleq \|T\| \cdot \|T^{-1}\|$ for some norm. Obviously, $\kappa(T) \leqslant \kappa(T)\kappa(B)\kappa(C)$, with $\leqq$ replaceable by $\ll$ if B and/or C is ill-conditioned.

Note that $\kappa(T)$ behaves like $(\min|\Delta_{iijj}|)^{-1}$ (cf. [9]).

3. THE NUMERICAL ALGORITHM

To avoid using complex arithmetic, the triangular real Schur forms of (A, C) and (D, B) will be used in Equation (8) instead. Let $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ be partitioned as

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & & & \bigcirc \\ \tilde{A}_{21} & \tilde{A}_{22} & & \\ \vdots & \vdots & \ddots & \\ \tilde{A}_{p1} & \tilde{A}_{p2} & \cdots & \tilde{A}_{pp} \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} \tilde{B}_{11} & \tilde{B}_{12} & \cdots & \tilde{B}_{1q} \\ & \tilde{B}_{22} & \cdots & \tilde{B}_{2q} \\ & & \ddots & \vdots \\ \bigcirc & & & \tilde{B}_{qq} \end{bmatrix},$$

$$\tilde{C} = \begin{bmatrix} \tilde{C}_{11} & & \bigcirc & \\ \tilde{C}_{21} & \tilde{C}_{22} & & \\ \vdots & \vdots & \ddots & \\ \tilde{C}_{p1} & \tilde{C}_{p2} & \cdots & \tilde{C}_{pp} \end{bmatrix}, \qquad \tilde{D} = \begin{bmatrix} \tilde{D}_{11} & \tilde{D}_{12} & \cdots & \tilde{D}_{1q} \\ & \tilde{D}_{22} & \cdots & \tilde{D}_{2q} \\ & & \ddots & \vdots \\ \bigcirc & & & \tilde{D}_{qq} \end{bmatrix},$$

with $\tilde{E}$ and $\tilde{X}$ conformally partitioned. Denoting $\tilde{A}_{ij}\tilde{X}_{jk}\tilde{B}_{k1} - \tilde{C}_{ij}\tilde{X}_{jk}\tilde{D}_{k1}$ by $T_{i1}(\tilde{X}_{jk})$, Equation (8) can now be written as

$$T_{ij}\left(\tilde{X}_{ij}\right) = \tilde{E}_{ij} - \sum_{\substack{k=1 \\ (k,l)\neq(i,j)}}^{i} \sum_{l=1}^{j} T_{ij}\left(\tilde{X}_{kl}\right), \tag{13}$$

with $i = 1,\dots,p$ and $j = 1,\dots,q$.

Again, if $\tilde{X}_{ij}$ are calculated in a rowwise fashion (or columnwise if preferred), the terms on the RHS of the equation (13) are all known. Equation (13) is then a linear equation involving components of $\tilde{X}_{ij}$ and is at most 4×4. It can be solved by the Kronecker tensor product and Gaussian elimination approach for each i and j, as in Equation (3). [Equation (13) can of course be scalar.]

After the solution of Equation (13) for all the i and j, one can then retrieve $X = P_2\tilde{X}Q_1$ from $\tilde{X}$.

The numerical algorithm can then be summarized as follows:

Algorithm 2.

Step 1. Transform (A, C) by the QZ algorithm to *lower*-triangular real Schur form.
Stop if $A - \lambda C$ is a (nearly) singular pencil.

Step 2. Transform (D, B) by the QZ algorithm to *upper*-triangular real Schur form.
Stop if $D - \lambda B$ is a (nearly) singular pencil.

Step 3. Consider cond(T), and stop if it is too large.

Step 4. Transform E to $\tilde{E}$.

Step 5. For $i = 1,\dots,p$ and $j = 1,\dots,q$, solve rowwise for $\tilde{X}_{ij}$ through the 4×4 or scalar system in Equation (13).

Step 6. Retrieve $X = P_2\tilde{X}Q_1$.

As was pointed out by the referees, the inspection of cond(T_{ij}) will be necessary but not sufficient to ensure the well-conditioning of T in step 3. One may try to estimate cond(T), using techniques generalizing those in LINEPACK [3]. In terms of operation counts, the estimation will be equivalent to solving Equation (1) with another RHS.

It may also be interesting to consider the scaling problem for Equation (1).

Note that a backward error analysis is not yet available for Algorithm 2, although the individual components of the algorithm are numerically stable.

The refinement idea [1, 12] can easily be implemented.

Note that the tricks and remarks in [1] (e.g. modifications for symmetric matrices) mostly apply to Algorithm 2.

4. OPERATION COUNTS

In this section, an operation count is presented for Algorithm 2. A count for Epton's method [4] is also presented as a comparison.

In Epton's method, (D, B) is transformed to upper-triangular Schur form (which is in general complex) and (A, C) to upper- (lower-) triangular Hessenberg form. Using similar notation to that in Equation (8), one can solve for $\tilde{x}_j$, the jth column of $\tilde{X}$, through

$$(\beta_{jj}A - \delta_{jj}C)\tilde{x}_j = \tilde{e}_j - \sum_{i<j}(\beta_{ij}A - \delta_{ij}C)\tilde{x}_i, \tag{14}$$

$j = 1, \ldots, n$, where $\tilde{e}_j$ is the jth column of $\tilde{E}$.

The matrix $\beta_{jj}A - \delta_{jj}C$ is Hessenberg, and Equation (14) can be solved efficiently. Note that the method relies on the strict upper-triangular features of D and B, and complex arithmetic is unavoidable unless modifications making use of the real Schur form are carried out.

For a system (1) with N different right-hand sides E, Algorithm 2 in Section 3 requires approximately

$$c_1 \cong 15(m^3 + n^3) + N\cdot\{18(m^3 + n^3) + 4(mn^2 + nm^2)\} \text{ real flops,}$$

with

$$c_1 \cong (30 + 44N)n^3 \qquad \text{when} \quad m \cong n$$

and

$$c_1 \cong (15 + 18N)m^3 \qquad \text{when} \quad m \gg n.$$

c_1 is obtained assuming that only two interations are required for each eigenvalue block in the QZ algorithm in steps 1 and 2 of Algorithm 2, and all the systems in Equation (13) are 4×4.

Similarly, for Epton's algorithm, one has

$$c_2 \cong 5m^3 + 15n^3 + 4nm^2 + N\cdot\left\{3\tfrac{5}{6}m^3 + 18n^3 + 9nm^2 + 3mn^2\right\} \text{ real flops},$$

with

$$c_2 \cong \left(24 + 33\tfrac{5}{6}N\right)n^3 \qquad \text{when} \quad m \cong n$$

and

$$c_2 \cong \left(5 + 3\tfrac{5}{6}N\right)m^3 \qquad \text{when} \quad m \gg n.$$

Obviously, $c_1 > c_2$, especially when $m \gg n$. Note that c_1 and c_2 are dominated by the transformations of the matrices A, B, C, D, and E to various standard canonical forms.

As a conclusion, Epton's method may be preferable, especially when $m \gg n$. Otherwise, the method in Section 3 may be preferred, especially when $m \cong n$, or when (A, C) and (D, B) are already in triangular real Schur form, e.g. when one is also interested in the spectra of (A, C) and (D, B).

5. LEAST-SQUARES SOLUTIONS

Consider the generalization of Equation (1) where $\rho(A, C) \cap \rho(D, B) \neq \varnothing$ and the matrix pencils $A - \lambda C$ and $D - \lambda B$ are allowed to be singular, or indeed rectangular. (See [5], [10].) One can then analyse the structures of (A, C) and (D, B) by using the Van Dooren algorithm [10]. In the transformed form, Equation (1) can then be written as

$$\begin{bmatrix} \tilde{A}_R & 0 & 0 \\ \tilde{A}_{21} & \tilde{A}_I & 0 \\ \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_S \end{bmatrix} \tilde{X} \begin{bmatrix} \tilde{B}_R & \tilde{B}_{12} & \tilde{B}_{13} \\ 0 & \tilde{B}_I & \tilde{B}_{23} \\ 0 & 0 & \tilde{B}_S \end{bmatrix}$$
$$- \begin{bmatrix} \tilde{C}_R & 0 & 0 \\ \tilde{C}_{21} & \tilde{C}_I & 0 \\ \tilde{C}_{31} & \tilde{C}_{32} & \tilde{C}_S \end{bmatrix} \tilde{X} \begin{bmatrix} \tilde{D}_R & \tilde{D}_{12} & \tilde{D}_{13} \\ 0 & \tilde{D}_I & \tilde{D}_{23} \\ 0 & 0 & \tilde{D}_S \end{bmatrix}$$
$$= \begin{bmatrix} \overline{E}_{11} & \tilde{E}_{12} & \tilde{E}_{13} \\ \tilde{E}_{21} & \tilde{E}_{22} & \tilde{E}_{23} \\ \tilde{E}_{31} & \tilde{E}_{32} & \tilde{E}_{33} \end{bmatrix}, \tag{15}$$

with $\tilde{X}$ partitioned similarly to $\tilde{E}$. The suffixes R and I represent the regular part of the matrix pencils, and S the singular part. The regular part is further divided into two disjoint parts, with I denoting the part with intersecting spectra.

Equation (15) can be broken up into

$$\begin{aligned}
&\tilde{A}_R\tilde{X}_{11}\tilde{B}_R - \tilde{C}_R\tilde{X}_{11}\tilde{D}_R = \tilde{E}_{11},\\
&\tilde{A}_R\tilde{X}_{12}\tilde{B}_I - \tilde{C}_R\tilde{X}_{12}\tilde{D}_I = \tilde{E}_{12} - \tilde{A}_R\tilde{X}_{11}\tilde{B}_{12} + \tilde{C}_R\tilde{X}_{11}\tilde{D}_{12},\\
&\tilde{A}_I\tilde{X}_{21}\tilde{B}_R - \tilde{C}_I\tilde{X}_{21}\tilde{D}_R = \tilde{E}_{21} - \tilde{A}_{21}\tilde{X}_{11}\tilde{B}_R + \tilde{C}_{21}\tilde{X}_{11}\tilde{D}_R,
\end{aligned} \tag{16}$$

together with other equations which cannot be solved by Algorithm 2 in the usual non-least-squares sense. $\tilde{X}_{11}, \tilde{X}_{12}, \tilde{X}_{21}$ can then be found using Algorithm 2, and substituted back into the other equations. They can then be written down in Kronecker tensor form and solved in the least-squares sense, e.g. using the QR decomposition. The idea should be viable if the dimensions of $\tilde{A}_I$, $\tilde{A}_S$, $\tilde{B}_I$, and $\tilde{B}_S$, i.e. the intersecting and singular parts of the matrix pencils, are small.

Note that if the matrix pencils are regular, the only other equations will be

$$\tilde{A}_I\tilde{X}_{22}\tilde{B}_I - \tilde{C}_I\tilde{X}_{22}\tilde{D}_I = \tilde{E}_{22} + \text{terms involving } \tilde{X}_{11},\ \tilde{X}_{12}, \text{ and } \tilde{X}_{21}. \tag{17}$$

The other equations can be both under- and overdetermined at the same time, e.g. when

$$A - \lambda C = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = D = 1 \quad \text{(scalar)};$$

then $A - \lambda C$ is purely singular and is in Kronecker canonical form [5]. Equation (1) is then equivalent to

$$x_1 + x_2 = e_1, \tag{18a}$$

$$x_3 = e_2, \tag{18b}$$

$$x_3 = e_3. \tag{18c}$$

Obviously, Equation (18a) is underdetermined for x_1 and x_2, and Equations (18b) and (18c) overdetermined for x_3.

Finally, an eigenvalue problem always has a parallel linear system of equations, and it is interesting to see the way the "singularity" in singular matrix pencils manifests itself through the under- and overdetermined set of linear equations in the form of Equation (1).

6. THE SIMULTANEOUS EQUATIONS $(YA - DZ,\ YC - BZ) = (E, F)$

In [9], Stewart introduced the operator T defined by

$$T[(Y, Z)] \triangleq (YA - DZ, YC - BZ) = (E, F), \tag{19}$$

and proved that, for systems which satisfied condition (i) in Theorem 1, T is invertible if and only if condition (ii) in Theorem 1 holds. (One can prove a slightly stronger result, as in Theorem 4 below.) Obviously, the operators in Equations (12) and (19) are closely related.

Assuming that $A - \lambda C$ and $D - \lambda B$ are invertible for some $\lambda \in \mathbb{R}$. Equation (19) can then be written as

$$Y = (D - \lambda B)X + E - \lambda F, \tag{20a}$$

$$Z = X(A - \lambda C), \tag{20b}$$

$$DXC - BXA = F + (\lambda F - E)(A - \lambda C)^{-1}C. \tag{20c}$$

Note that Equation (20c) is in the form of Equation (1), and Y and Z can be evaluated through Equations (20a) and (20b) after X has been obtained by solving equation (20c). However, that would be unwise, as the inversion of $A - \lambda C$ is involved.

Starting from Equation (1), define $Y = (A - \lambda C)X$ and $Z = X(D - \lambda B)$. Equation (1) can easily be proved to be of the form

$$(YB - CZ, YD - AZ) = (E, \lambda E), \tag{21}$$

which is in the form of Equation (19). Again, solving Equation (1) through Equation (21) is not advisable, as the inversion of the matrix $A - \lambda C$ or $D - \lambda B$ is involved.

As a result, it is proved that the solutions of Equations (1) and (19) are equivalent for systems with a unique solution (or satisfying the conditions in Theorem 1).

We are now ready to prove the following theorem:

THEOREM 4. *The matrix equation* (19) *has a unique solution if and only if conditions* (i) *and* (ii) *of Theorem* 1 *are satisfied.*

Proof. The "if" part has been proved by the above argument. (It can also be proved by a similar argument to that in the proof of Theorem 1, after transforming Equation (19) into a Sylvester equation. See also [9].) The "only if" part can be proved as follows: Consider the equation

$$Y(\lambda_1 A - \lambda_2 C) - (\lambda_1 D - \lambda_2 B)Z = \lambda_1 E - \lambda_2 F \tag{22}$$

for some real λ_1 and λ_2 which are not both zero. Equation (22) can then replace one of the two equations in (19) and still leave an equivalent set of equations. If either or both of the conditions (i) and (ii) are violated, the homogeneous equation related to Equation (22) will be satisfied by $Y = y_1 y_2^H$ and $Z = z_1 z_1^H$, with $y_2^H(\lambda_1 A - \lambda_2 C) = 0$ and $(\lambda_1 D - \lambda_2 B)z_1 = 0$ for some chosen λ_i's. Let the remaining equation be, without loss of generality, $YA - DZ = E$, with its related homogeneous equation satisfied by choosing $y_1 = Dz_1$ and $z_2 = A^H y_2$. Thus a solution of Equation (19) cannot be unique, if it exists. ■

One can generalize the concepts of "diff" in [9] and relate it to $\|T^{-1}\|$, for the operator T in Equation (12).

A similar procedure to that in Algorithm 2 for Equation (19) is as follows:

Algorithm 3.

Step 1. Transform (A, C) by the QZ algorithm to *upper*-triangular real Shur form.
Stop if $A - \lambda C$ is a (nearly) singular pencil.

Step 2. Transform (D, B) by the QZ algorithm to *lower*-triangular real Schur form.
Stop if $D - \lambda B$ is a (nearly) singular pencil.

Step 3. Consider $\operatorname{cond}(T)$, and stop if it is too large.

Step 4. Transform (E, F) to $(\tilde{E}, \tilde{F})$.

Step 5. Equation (19) is then equivalent to

$$\tilde{Y}_{ij}\tilde{A}_{jj} - \tilde{D}_{ii}\tilde{Z}_{ij} = \tilde{E}_{ij} - \sum_{k=1}^{j-1} \tilde{Y}_{ik}\tilde{A}_{kj} + \sum_{l=1}^{i-1} \tilde{D}_{il}\tilde{Z}_{lj},$$

$$\tilde{Y}_{ij}\tilde{C}_{jj} - \tilde{B}_{ii}\tilde{Z}_{ij} = \tilde{F}_{ij} - \sum_{k=1}^{j-1} \tilde{Y}_{ik}\tilde{C}_{kj} + \sum_{l=1}^{i-1} \tilde{B}_{il}\tilde{Z}_{lj}$$

for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. If $\tilde{Y}_{ij}$ and $\tilde{Z}_{ij}$ are calculated in a rowwise (columnwise, if preferred) fashion, the RHS will contain only known quantities and one will have to solve an 8×8 or 2×2 system for each i and j, for the components of $\tilde{Y}_{ij}$ and $\tilde{Z}_{ij}$.

Step 6. Retrieve Y and Z from $\tilde{Y}$ and $\tilde{Z}$.

The comments on ill-conditioning in Section 3 also apply here.

Again, modifications for symmetric matrices are possible, as in [1], to improve efficiency.

Finally, the equivalence between Equations (1) and (19) breaks down for systems involving nonunique solutions or singular matrix pencils. The solution of such equations in the least-squares sense is feasible, analogously to the techniques discussed in Section 5.

7. CONCLUSIONS

The necessary and sufficient conditions for the existence and uniqueness of the solution of the matrix Equation (1) has been presented. A numerical algorithm is proposed. An operation count has been given and compared with that of Epton's method [4]. The possibility of solving a general rectangular system in the form of Equation (1) in the least square sense has been briefly discussed.

Equation (1) has been proved to be equivalent to Equation (19), when a unique solution exists. A numerical algorithm for the solution of Equation (19) has been proposed.

Finally, note that Theorem 1 and Algorithm 2 can be generalized with ease for the equation

$$\sum_{i=1}^{p} f_{1i}(A)Xf_{2i}(B)+\sum_{i=1}^{q} f_{3i}(C)Xf_{4i}(D)=C,$$

if the functions $f_{ij}(M)$ preserve the triangular structure of the matrix M (e.g. polynomials, exponential e^M). The conditions for the solvability will then be

(i) $\mathscr{D}_1 \triangleq \det[\Sigma_{i=1}^{p} f_{1i}(\lambda_1 A)+\Sigma_{i=1}^{q} f_{3i}(\lambda_3 C)]$ and $\mathscr{D}_2 \triangleq \det[\Sigma_{i=1}^{q} f_{4i}(\lambda_4 D)+\Sigma_{i=1}^{p} f_{2i}(\lambda_2 B)]$ are not identically zero, and

(ii) $\rho_1 \cap \rho_2 = \varnothing$, with

$$\rho_1 \triangleq \{(\lambda_1,\lambda_3) : \mathscr{D}_1 = 0\} \quad \text{and} \quad \rho_2 \triangleq \{(\lambda_4,\lambda_2) : \mathscr{D}_2 = 0\}.$$

This paper was written while the author was supported by the SERC of U.K., contract number GR/C/95190.

Thanks are also due to the referees for several valuable comments.

REFERENCES

1 R. H. Bartels and G. W. Stewart, Solution of the matrix equation $AX + XB = C$, *Comm. ACM* 15:820–826 (Sept. 1972).
2 K.-W. E. Chu, Exclusion Theorems for The Generalized Eigenvalue Problem, Numer. Anal. Rept. NA/11/85, Dept. of Mathematics, Univ. of Reading, Reading, U.K., 1985 (to appear in *SIAM J. Numer. Anal.*).
3 A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, An estimate for the condition number of a matrix, *SIAM J. Numer. Anal.* 16:368–375 (Apr. 1979).
4 M. A. Epton, Methods for the solution of $AXD - BXC = E$ and its application in the numerical solution of implicit ordinary differential equations, *BIT* 20:341–345 (1980).
5 F. R. Gantmacher, *The Theory of Matrices, II*, Chelsea, New York, 1960.
6 G. H. Golub, S. Nash, and C. Van Loan, A Hessenberg-Schur Method for the Matrix Problem $AX + XB = C$, *IEEE Trans. Automat. Control* AC-24:909–913 (1979).
7 C. B. Moler and G. W. Stewart, An algorithm for the generalized matrix eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 10:241–256 (*Apr.* 1973).
8 M. Z. Nashed, *Generalized Inverses and Applications*, Academic, New York, 1976.
9 G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 9:669–686 (Dec. 1972).
10 P. Van Dooren, The computation of Kronecker's canonical form of a singular pencil, *Linear Algebra Appl.* 27:103–140 (1979).
11 R. C. Ward, The combination shift QZ algorithm, *SIAM J. Numer. Anal.* 12:835–852 (Dec. 1975).
12 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 2 August 1985; revised 24 July 1986