

The Solution of The Matrix Equations

$AXB - CXD = E$  and  $(YA - DZ, YC - BZ) = (E, F)$

by

Dr. King-wah Eric Chu

NUMERICAL ANALYSIS REPORT No. NA/10/85

Running Short Title:-

The Solution of The Matrix Equations

## Abstract

The conditions for the existence of a unique solution of the matrix equation  $AXB - CXD = E$  is proved to be (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 its solution is proposed. The possibility of a least square type solution is briefly discussed.

The set of equations  $(YA-DZ, YC-BZ) = (E,F)$  is proved to be equivalent to the afore-mentioned equation, and its solution is also investigated. A numerical algorithm is proposed.

## Acknowledgement

The author is employed under a SERC research contract.

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 of  $X$ :

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

By using the Kronecker tensor product, equation (2) (and thus (1)) can be written as

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

where  $v(X) = (x_1^T, x_2^T, \dots, x_m^T)^T$ , with  $x_i^T$  the  $i$ -th rows 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], it 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  $O(m^3 n^3)$  flops. The operation count is obviously unacceptable when  $m \cong n$ .

In [4], equation (1) has been solved directly by transforming  $(A, C)$  to upper-triangular-Hessenberg form and  $(D, B)$  to upper-triangular-Schur form. (See section 4 for more details.) The operation count is  $O(m^3) + O(n^3)$  flops. However, no condition for the solvability of

equation (1) is presented and no test on the solvability is carried out in the algorithm before the equation is solved. In addition, the matrix pencils  $(A-\lambda C)$  and  $(D-\lambda B)$  are assumed to be regular, but no test is carried out on  $(A-\lambda C)$ . Note that a matrix pencil can be regular, but nearly singular, especially when round-off errors are taken into account.

In this paper, a set of necessary and sufficient conditions for the existence and uniqueness of the solution of equation (1) is presented. A stable numerical algorithm for the solution is proposed. The algorithm is slightly less efficient than Epton's [4] in terms of operation counts, especially when  $m \gg n$ , but equations involving singular pencils, non-existent or non-unique solutions will be detected by the algorithm. Possibilities of solving equation (1) in the least square sense is discussed briefly in section 5.

In section 6, equation (1) is proved to be equivalent to the set of equations

$$\begin{cases} YA - DZ = E \\ YC - BZ = F \end{cases} \quad (19)$$

and a stable numerical algorithm is proposed for the solution of equation (19).

Note that the equation (1) is a generalization of the Sylvester equation  $AX - XD = E$ , discussed by Bartels and Stewart [1], and 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 [3]. Other applications of the equation (1) can be found in [4] and the references therein.

2. The Solvability of Equation (1).

Consider the generalized eigenvalue problem [3][5][7][9][11][12]

$$Ax = \lambda Cx \quad (4)$$

in the more sensible and convenient form

$$\gamma Ax = \alpha Cx, \quad (5)$$

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  $\lambda$  will now be treated similarly as  $(\alpha, \gamma) = (0, \gamma)$  or  $(\alpha, 0)$ . From equations (4) and (5), one has

$$\lambda = \alpha/\gamma, \quad (6)$$

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

Consider a regular pencil  $(A-\lambda C)$ . In general, there exists unitary matrices  $P_1$  and  $P_2$ , through the QZ decomposition [7][11], such that  $P_1 A P_2^{\Delta}(\alpha_{ij})$  and  $P_1 C P_2^{\Delta}(\gamma_{ij})$  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 exists unitary matrices  $Q_1$  and  $Q_2$  such that  $Q_1 D Q_2^{\Delta}(\delta_{ij})$  and  $Q_1 B Q_2^{\Delta}(\beta_{ij})$  are both upper triangular, with  $(D-\lambda B)$  a regular pencil. (c.f. [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

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

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

Equation (1) has now been transformed to

$$P_1 A P_2^H \cdot P_2^H X Q_1^H \cdot Q_1 B Q_2 - P_1 C P_2^H \cdot P_2^H X Q_1^H \cdot Q_1 D Q_2 = P_1 E Q_2$$

$$\Leftrightarrow \tilde{A} \tilde{X} \tilde{B} = \tilde{C} \tilde{X} \tilde{D} = \tilde{E} \quad (\tilde{E}(\epsilon_{ij})) \quad (8)$$

$$(\Delta \tilde{X} \nabla = \tilde{X} \nabla = \tilde{X})$$

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  $\Delta_{ijkl}$  denoting  $(\alpha_{ij} \beta_{kl} - \gamma_{ij} \delta_{kl})$ ,

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

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

and for a general  $(i,j)$ ,

$$\Delta_{iijj} \cdot \tilde{x}_{ij} = \epsilon_{ij} - \sum_{k=1}^{i-1} \sum_{l=1}^{j-1} (\Delta_{iklj} \cdot \tilde{x}_{kl}) \quad (9)$$

It is obvious from equation (9) that equation (8), and thus (1), can be solved for a unique  $X$ , if and only if  $\Delta_{iijj} \neq 0, \forall i,j$ ;

$$\Leftrightarrow \rho(A,C) \cap \rho(D,B) = \emptyset.$$

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)  $\rho(A,C) \cap \rho(D,B) = \emptyset$ . (Recall the equivalence classes defined by equation (7).)

(Proof) Consider the equations

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

and

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

for some real  $\lambda_1$  and  $\lambda_2$  which are not both zero.

One of the equations in (10) is equivalent to equation (1).

(c.f. [4].)

If the conditions (i) and (ii) are satisfied,  $\lambda_i$ 's can be found so that the matrices involving the  $\lambda_i$ 's are non-singular, thus solving equation (10a) or (10b) is equivalent to solving an equivalent Sylvester equation, which yields a unique solution.

If any one of the conditions (i) and (ii) (or both) is (are) violated, some  $\lambda_i$ '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  $z \neq 0$  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 non-zero constant  $c$ , will be a non-trivial 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 reduces to  $\rho(A) \cap \rho(D) = \emptyset$ . (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) will then be detected by inspecting the spectra  $\rho(A,C)$  and  $\rho(D,B)$ ,

after the QZ processes have been performed in equation (8). Singular or nearly singular matrix pencils can also be detected the same way.

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

$$C^{-1}AX - XDB^{-1} = C^{-1}EB^{-1}, \quad (11)$$

one should not solve equation (1) in the form of equation (11).

Denote the equation (1), using the operator T, as

$$T(X) \triangleq AXB - CXD = E, \quad (12)$$

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

Note that  $\kappa(T)$  behaves like  $(\min |\Delta_{ijjj}|)^{-1}$ . (c.f. [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

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

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

with  $\tilde{E}$  and  $\tilde{X}$  conformally partitioned. Denote

$(\tilde{A}_{ij}\tilde{X}_{jk}\tilde{B}_{kl} - \tilde{C}_{ij}\tilde{X}_{jk}\tilde{D}_{kl})$  by  $T_{ij}(\tilde{X}_{jk})$ , equation (8) can now be written as

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

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

Again, if  $\tilde{X}_{ij}$  are calculated in a row-wise fashion, (or column-wise 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  $4 \times 4$ . 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. Calculate the eigenvalues  $(\alpha, \gamma)$  and  $(\delta, \beta)$  and stop if condition (ii) is (nearly) violated. Check  $\min |\alpha_{ii} \beta_{jj} - \gamma_{ii} \delta_{jj}|$  for conditioning.
- Step 4. Transform  $E$  to  $\tilde{E}$ .
- Step 5. For  $i = 1, \dots, p$  and  $j = 1, \dots, q$ ; solve row-wise for  $\tilde{X}_{ij}$  through the  $4 \times 4$  or scalar system in equation (13).
- Step 6. Retrieve  $X = P_2 \tilde{X} Q_1$ .

The method should be numerically stable, in view of the stable numerical behaviour of the individual component algorithms used. In addition, the stability and round-off error analysis in [6] can be modified to cope with the above algorithm.

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 the afore-mentioned Algorithm 2.

#### 4. Operation Counts.

In this section, an operation count is presented for the 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 notations as in equation (8), one can solve for  $\tilde{x}_j$ , the  $j$ -th column of  $\tilde{X}$ , through

$$(\beta_{jj}A - \delta_{jj}C)x_j = \tilde{e}_j - \sum_{i < j} (\beta_{ij}A - \delta_{ij}C)x_i, \quad (14)$$

$j = 1, \dots, n$ ; where  $\tilde{e}_j$  is the  $j$ -th 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 some complex arithmetic is unavoidable.

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

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

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

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

$c_1$  is obtained assuming that only 2 iterations 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  $4 \times 4$ .

Similarly, for Epton's algorithm, one has

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

$$\text{with } c_2 \cong (24 + 33\frac{5}{6}N)n^3, \text{ when } m \cong n,$$

$$\text{and } c_2 \cong (5 + 3\frac{5}{6}N)m^3, \text{ when } m \gg n.$$

Obviously,  $c_1 > c_2$ , especially when  $m \gg n$ . However, ill-conditioning of equation (1) can only be detached through the LU-decompositions of the matrices  $(\beta_{jj}A - \delta_{jj}C)$  in equation (14), and it may well be after a lot of work has been done. 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 should be used if one is sure about the solvability and conditioning of the equation (1), especially when  $m \gg n$ . Otherwise, the method in section 3 should be preferred, especially when  $m \cong n$ . In addition,  $c_1$  will be less than  $c_2$  if  $(A, C)$  is already in lower-triangular-real-Schur form, e.g. when one is also interested in the spectra of  $(A, C)$  and  $(D, B)$ .

5. Least Square Solutions.

Consider the generalization of equation (1), where  $\rho(A,C) \cap \rho(D,B) \neq \emptyset$  and the matrix pencils  $(A-\lambda C)$  and  $(D-\lambda B)$  are allowed to 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

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

with  $\tilde{X}$  similarly partitioned as  $\tilde{E}$ .

The suffices 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

$$\begin{cases} \tilde{A}_R \tilde{X}_{11} \tilde{B}_R - \tilde{C}_R \tilde{X}_{11} \tilde{D}_R = \tilde{E}_{11}, \\ \tilde{A}_R \tilde{X}_{12} \tilde{B}_I - \tilde{C}_R \tilde{X}_{12} \tilde{D}_I = \tilde{E}_{12} - \tilde{A}_R \tilde{X}_{11} \tilde{B}_{12} + \tilde{C}_R \tilde{X}_{11} \tilde{D}_{12}, \\ \tilde{A}_I \tilde{X}_{21} \tilde{B}_R - \tilde{C}_I \tilde{X}_{21} \tilde{D}_R = \tilde{E}_{21} - \tilde{A}_{21} \tilde{X}_{11} \tilde{B}_R + \tilde{C}_{21} \tilde{X}_{11} \tilde{D}_R, \end{cases} \quad (16)$$

with "other equations" which cannot be solved by Algorithm 2 in the usual non-least-square sense.  $\tilde{X}_{11}, \tilde{X}_{12}, \tilde{X}_{21}$  can then be solved 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 square 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

$$\begin{aligned} & \tilde{A}_I \tilde{X}_{22} \tilde{B}_I - \tilde{C}_I \tilde{X}_{22} \tilde{D}_I \\ = & \tilde{E}_{22} + \text{terms involving } \tilde{X}_{11}, \tilde{X}_{12} \text{ and } \tilde{X}_{21}. \end{aligned} \quad (17)$$

The "other equations" can be both under- and over-determined at the same time, e.g. when  $(A-\lambda C) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda \end{pmatrix}$ ,  $B = D = 1$  (scalar);

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

Equation (1) is then equivalent to

$$\begin{cases} x_1 + x_2 = e_1 & (18a) \\ x_3 = e_2 & (18b) \\ x_3 = e_3. & (18c) \end{cases}$$

Obviously, equation (18a) is under-determined for  $x_1$  and  $x_2$ , with equations (18b) and (18c) over-determined 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 over-determined 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  $T$ , where

$$T[(Y,Z)]^A (YA-DZ, YC-BZ) = (E,F), \quad (19)$$

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

$$\begin{cases} Y = (D-\lambda B)X + E - \lambda F, & (20a) \\ Z = X(A-\lambda C), & (20b) \\ DXC - BXA = F + (\lambda F - E)(A-\lambda C)^{-1}C, & (20c) \end{cases}$$

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, it will 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

$$(YB-CZ, YD-AZ) = (E, \lambda E), \quad (21)$$

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 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 as 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

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

for some real  $\lambda_1$  and  $\lambda_2$  which are not both zero. Equation (22) can then replace one of the two equations in (19) and still leaves an equivalent set of equations. If any one of the conditions (i) and (ii) (or both) is (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  $\lambda_i$ '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 related it to  $\|T^{-1}\|$ , for the operator  $T$  in equation (12).

A similar procedure as in Algorithm 2 for equation (19) is as follows:

Algorithm 3.

Step 1. Transform  $(A,C)$  by the QZ algorithm to upper-triangular-real-Schur 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. Calculate the eigenvalues  $(\alpha,\gamma)$  and  $(\delta,\beta)$  and stop if condition (ii) is violated. Check  $\min|\alpha_{ii}\beta_{jj} - \gamma_{ii}\delta_{jj}|$  for conditioning.

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

Step 5. Equation (19) is then equivalent to

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

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

If  $\tilde{Y}_{ij}$  and  $\tilde{Z}_{ij}$  are calculated in a row-wise (column-wise, if preferred) fashion, the RHS will contain only known quantities and one will have to solve an  $8 \times 8$  or  $2 \times 2$  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 above Algorithm 3 is obviously numerically stable.

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 non-unique solutions or singular matrix pencils. The solution of such equations in the least square sense is feasible, analogous 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) is presented. A stable numerical algorithm is proposed. An operation count is given and compared to 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 is briefly discussed.

The equation (1) is then proved to be equivalent to equation (19), when a unique solution exists. A stable numerical algorithm for the solution of equation (19) is 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)  $\mathcal{D}_1 \triangleq \det \left[ \sum_{i=1}^p f_{1i}(\lambda_1 A) + \sum_{i=1}^q f_{3i}(\lambda_3 C) \right]$  and  $\mathcal{D}_2 \triangleq \det \left[ \sum_{i=1}^q f_{4i}(\lambda_4 D) + \sum_{i=1}^p f_{2i}(\lambda_2 B) \right]$  are not identically zero, and (ii)  $\rho_1 \cap \rho_2 = \emptyset$ , with

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

References.

- [1] R. H. Bartels & G. W. Stewart. Solution of The Matrix Equation  $AX + XB = C$ . ACM 15: 214-235 (Sept. 1972).
- [3] K. -W. E. Chu. Exclusion Theorems for The Generalized Eigenvalue Problem. Numer. Anal. Rpt., NA/11/85, Univ. of Reading, Dept. of Maths., Reading, U.K., 1985.
- [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 & C. Van Loan. A Hessenberg-Schur Method for The Matrix Problem  $AX + XB = C$ . IEEE Trans. Auto. Cont. AC-24: 909-913 (1979).
- [7] C. B. Moler & 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 Press, 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. Lin. Alg. Applic. 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.