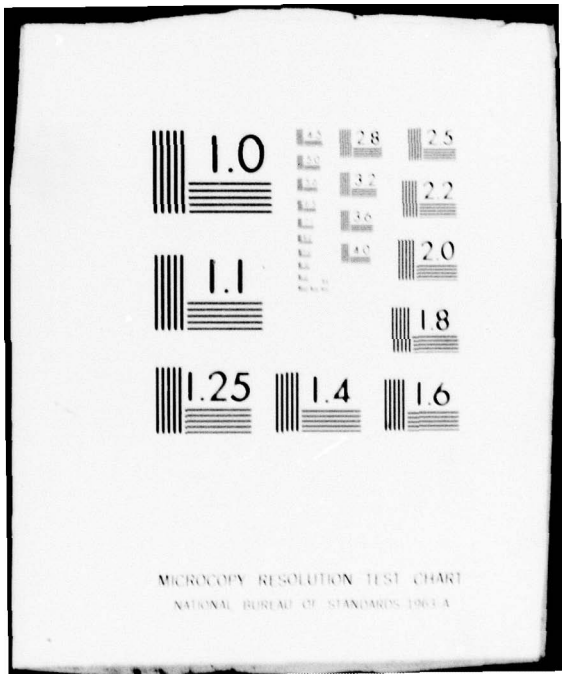
AD-A080 216 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOO--ETC F/G 12/1
SOLUTIONS OF THE MATRIX EQUATION AXB + CXD = E.(U)
DEC 79 T S POMYKALSKI
UNCLASSIFIED AFIT/GOR/MA/79D-6 NL

1 OF 2
AD A080216


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

UNITED STATES AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

(6) SOLUTIONS OF THE MATRIX EQUATION
AXB + CXD = E.

(9) Master's Thesis,

GOR/MA/79D-6

(10) Thomas S. Pomykalski
Captain USAF

(14) AFIT/GOR/MA/79D-6

(12) 106

(11) Dec 79

AFIT/GOR/MA/79D-6

SOLUTIONS OF THE MATRIX EQUATION

AXB + CXD = E

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science

by

Thomas S. Pomykalski
Captain USAF

Graduate Operations Research

December 1979

Preface

My studies of optimization theory and econometrics prompted me to take a further look into improving solution techniques for matrix equations fundamental to those areas. This thesis is a result of that investigation. It would not have been possible without the assistance and guidance of Dr. John Jones Jr., whose imagination and insights are its true source.

I would particularly like to thank Dr. J. Cain of AFIT, Dr. D.W. Repperger of the Aerospace Medical Laboratory, and Dr. R. Craig of the Air Force Materials Laboratory for taking the time to review this work.

Mrs. Suzanne Weber, my typist, deserves my most heartfelt thanks for taking this jumbled mass of equations and turning it into a clear, easily-readable format. This was truly a monumental task.

My wife and children have endured the most throughout this program of study. Their patience and love is without bound, and much appreciated.

Contents

	Page
Preface	ii
Abstract	iv
I. Introduction	1
II. Solutions of the Matrix Equation AXB+CXD = E.	4
Necessary and Sufficient Conditions for the Solution of Eq (2.1)	5
Conditions for the Solution of the Matrix Equation $\sum_{i=1}^{n} A_i X B_i = C$	11
Spectral Decomposition of Rectangular Matrices	29
Solutions of the Matrix Equation AXB+CYD=E.	40
Solutions Generated With the Use of Tensor Analysis	55
III. Generalized Inverses and Linear Models	61
Generalized Inverses	62
Linear Modeling	82
IV. Conclusions and Recommendations	91
Bibliography	92
Vita	96

Abstract

Solutions of the general linear matrix equation $\sum_{i=1}^{n} A_i X B_i = C$ are obtained and presented in this thesis. Some special cases do arise like the Liapanov matrix equation. Necessary conditions and sufficient conditions are established for the solution of the general linear matrix equation. Other forms of solutions than those obtained through the use of similarity transformations that have been considered make use of the spectral decomposition of matrices and tensor products of matrices or Kronecker products. In considering the general linear matrix equation, linear matrix equations in which two different variables appear are also studied. Conditions for the existence of a solution for this type of equation are given. The theory of the generalized inverses of a matrix was used in obtaining a solution to the general linear matrix equation. More general forms of the solution are given and conditions under which these solutions exist have been established. Solutions to systems of matrix equations were also considered. As a by-product of this investigation, some aspects of the model reduction problem may be treated from the point of view of matrix equations. In particular, a new method of solution of the matrix equation $AXB + CXD = E$ which was recently considered by S.K. Mitra (Siam Journal of Applied Math, 32) and others was obtained. Applications of the results of this

work are of use in the estimation of variance and covariance components of linear models as treated by C.R. Rao.

SOLUTIONS OF THE MATRIX EQUATION

$$AXB + CXD = E$$

I. Introduction

Matrix equations are becoming a more significant part of the formulation, computation and solution of problems in the engineering and social sciences. The classic engineering investigations of the Liapanov equation

$$AX + XB = C \tag{1.1}$$

and the matrix Ricatti equation

$$XDX + AX + XB + C = 0 \tag{1.2}$$

are probably the most commonly known. Solutions to these equations and their more general forms, given by the equation

$$\sum_{i=1}^{n} A_i X B_i = C \tag{1.3}$$

are studied in this thesis.

Solutions to Eq (1.3) are found through the application of various methods of solution. Necessary and sufficient conditions are stated for the case in which $n = 2$ in Eq (1.3). These conditions can easily be extended to cover larger values of n. The conditions given are predominently based upon the use of similar matrices. With similarity shown, the concept

of pencils of matrices is used to generate the actual solution to the matrix equation.

More complex matrix equations are handled through the introduction of the concept of spectral decomposition. This is basically writing a matrix in terms of eigenvalues and idempotent matrices. After some of the theory of spectral decomposition is investigated, it is extended by introducing nilpotent matrices. The theory of solutions is developed for square matrices first, and then extended to include rectangular coefficient matrices and also rectangular variable matrices. Following this is a brief digression to consider equations of the form

$$AXB + CYD = E \tag{1.4}$$

Equations of the type in Eq (1.4) are studied from the point of view of being able to be rewritten into an equation of the type in Eq (1.3) with $n = 2$.

Two other methods of solution of the equation $AXB + CXD = E$ are considered. A brief look is taken at how generalized inverses can be used to solve equations. In doing so, it becomes apparent that more than one solution can exist for an equation. The theory of how these additional solutions are generated is developed in Chapter III. The last method of solution considered is through the use of Tensor Analysis.

The second half of this thesis develops the concepts of generalized inverses by extending the methods that are currently being employed to allow a wider class of solution

to be possible. This is done by extending the theorems that exist to include more arbitrary matrices. This then lets the user have some control over the results and more closely fit the solution to the problem at hand. Also in this section, applications of the theory are made to the area of model reduction.

The last portion of the thesis recapitulates some of the ideas of linear modeling and then applies the previous work to finding the variance and covariance matrices. Another application is to the concept of the reduced order filter.

II. Solutions of the Matrix Equation

$AXB + CXD = E$

In this chapter the results of Roth (Ref 38), Rosenblum (Ref 36), Mitra (Ref 28), Jones (Ref 14), and Lancaster (Ref 21) are extended by use of procedures taken from Rao and Mitra (Ref 35), Nering (Ref 29), and Browne (Ref 4).

Special cases of the matrix equation

$$AXB + CXD = E \tag{2.1}$$

occur if $B = C = I$ or $A = D = I$, where I denotes the identity matrix. Thus, Eq (2.1) reduces to the Liapanov equation. Much attention has been given to solutions of the Liapanov equation, particularly Gantmacher (Ref 10), Ma (Ref 26), Rosenblum (Ref 36), and Ziedan (Ref 46). Many others have also done work in this area. Leuthauser (Ref 23) had studied Eq (2.1) in the cases where the matrices are all taken to be square. The results presented in the remainder of this chapter mostly pertain to rectangular matrices. That is, in general the solution matrix X is an element of a class of matrices that are of dimension m by n. This will be denoted by $X_{(m,n)}$. The dimensions of the other matrices are: $A_{(p,m)}$, $B_{(n,q)}$, $C_{(p,m)}$, $D_{(n,q)}$, and $E_{(p,q)}$.

Solutions of Eq (2.1) and also of the more general matrix equation

$$\sum_{i=1}^{m} A\, XB = Q \tag{2.2}$$

will be found through use of matrix methods, tensor products, and spectral decomposition of matrices.

Necessary and Sufficient Conditions for the Solution of Eq (2.1)

Roth (Ref 38) developed his results for square matrices by using the concept of similar matrices. Since the equation now under consideration has rectangular component matrices, the restriction that the matrices be similar is relaxed and all that is required is that the matrices be equivalent.

Definition: A matrix B is said to be equivalent to a matrix A if there exist nonsingular matrices P and Q such that $B = PAQ$.

Theorem 2.1 (Necessary Condition): Let $X_{(m,n)}$ denote the matrix solution of $AXB + CXD = E$ where $A_{(p,m)}$, $B_{(n,q)}$, $C_{(p,m)}$, $D_{(n,q)}$, and $E_{(p,q)}$, then the fcllowing pair of matrices

$$\begin{bmatrix} A_{(p,m)} & E_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} \quad , \quad \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} \tag{2.3}$$

are equivalent.

Proof: To show that the above matrices are equivalent, there must exist two nonsingular matrices P and Q such that,

$$P_{(p+n,p+n)} \cdot \begin{bmatrix} A_{(p,m)} & E_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} \cdot Q_{(m+q,m+q)} = \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} \tag{2.4}$$

$$\text{Let } P = \begin{bmatrix} I_{(p,p)} & -CX_{(p,n)} \\ O_{(n,p)} & I_{(n,n)} \end{bmatrix} \quad \text{and } Q = \begin{bmatrix} I_{(m,m)} & -XB_{(m,q)} \\ O_{(q,m)} & I_{(q,q)} \end{bmatrix} \tag{2.5}$$

Then by substitution, the resulting equation is

$$\begin{bmatrix} I_{(p,p)} & -CX_{(p,m)} \\ O_{(n,p)} & I_{(n,n)} \end{bmatrix} \cdot \begin{bmatrix} A_{(p,m)} & E_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix}$$

$$\cdot \begin{bmatrix} I_{(m,m)} & -XB_{(m,q)} \\ O_{(q,m)} & I_{(q,q)} \end{bmatrix} = \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(m,n)} & D_{(n,q)} \end{bmatrix} \tag{2.6}$$

Multiplying the first two matrices on the left together yields:

$$\begin{bmatrix} A_{(p,m)} & E\text{-}CXD_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} \cdot \begin{bmatrix} I_{(m,m)} & -XB_{(m,q)} \\ O_{(q,m)} & I_{(q,q)} \end{bmatrix}$$

$$= \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(m,n)} & D_{(n,q)} \end{bmatrix} \tag{2.7}$$

Now, multiplying the remaining two matrices yields,

$$\begin{bmatrix} A_{(p,m)} & -AXB+E-CXD_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} = \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(m,n)} & D_{(n,q)} \end{bmatrix} \tag{2.8}$$

Since AXB+CXD=E, then -AXB+E-CXD=0 and Eq (2.8) becomes

$$\begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(n,m)} & D_{(n,q)} \end{bmatrix} = \begin{bmatrix} A_{(p,m)} & O_{(p,q)} \\ O_{(m,n)} & D_{(n,q)} \end{bmatrix} \tag{2.9}$$

which shows the matrices to be equivalent and completes the proof.

Theorem 1 shows that a necessary condition for a solution to exist is that the matrices A and D from Eq (2.1) can be written in block diagonal form. Similarly, the same result can be shown for the matrices B and C of Eq (2.1).

Theorem 2.2 (Necessary Condition): Let $X_{(m,n)}$ denote the matrix solution of AXB+CXD=E where $A_{(p,m)}$, $B_{(n,q)}$, $C_{(p,m)}$, $D_{(n,q)}$ and $E_{(p,q)}$, then the following pair of matrices are equivalent:

$$\begin{bmatrix} C_{(p,m)} & E_{(p,q)} \\ O_{(n,m)} & B_{(n,q)} \end{bmatrix} \quad , \quad \begin{bmatrix} C_{(p,m)} & O_{(p,q)} \\ O_{(n,m)} & B_{(n,q)} \end{bmatrix} \qquad (2.10)$$

Proof: The proof of Theorem 2.2 is identical to the proof of Theorem 2.1 with the choice of P and Q as follows:

$$\text{Let } P = \begin{bmatrix} I_{(p,p)} & -AX_{(p,n)} \\ O_{(n,p)} & I_{(n,n)} \end{bmatrix} \text{ and } Q = \begin{bmatrix} I_{(m,m)} & -XD_{(m,q)} \\ O_{(q,m)} & I_{(q,q)} \end{bmatrix} \qquad (2.11)$$

Two other necessary conditions arise if in Eq (2.1) the matrices A and C are of the same square dimension and the matrices B and D are also of the same square dimension but different from A and C. These conditions arise from the study of pencils.

Definition: Let A anc C be a pair of square matrices of the same order and let λ be an element of the complex numbers, then $A + \lambda C$ is called a pencil.

Definition: A pencil is considered to be regular if the matrices A and C are square and the determinant $|A + \lambda C|$

is not identically zero. Otherwise, the pencil is singular (Ref 28:823).

Theorem 2.3: If $A+\lambda C$ and $B+\lambda D$ are regular pencils, then Eq (2.1) has a solution if and only if the following pair of matrices

$$\begin{bmatrix} (C-eA)^{-1} & (C-eA)^{-1}E(B+eD)^{-1} \\ 0 & -D(B+eD)^{-1} \end{bmatrix}, \quad \begin{bmatrix} (C-eA)^{-1} & 0 \\ 0 & -D(B+eD)^{-1} \end{bmatrix} \tag{2.12}$$

are similar, where there exists an e, a complex scalar, such that $|-eA+C| \neq 0$ and $|B+eD| \neq 0$.

Proof: Since $A+\lambda C$ and $B+\lambda D$ are regular pencils, the determinants are $|A+\lambda C| \neq 0$ and $|B+\lambda D| \neq 0$. This implies that $|A+\lambda C|$ is a polynomial in λ and has at most n-zeros. Hence, there are at most m values of λ for which $|A+\lambda C|$ vanishes. This is similarly true for $B+\lambda D$. Choose e not equal to any value of λ for which $|A+\lambda C|$, $|B+\lambda D|$ vanishes (Ref 8:824). Thus e is a scalar such that $|-eA+C| \neq 0$ and $|B+eD| \neq 0$, which implies $(-eA+C)^{-1}$ and $(B+eD)^{-1}$ both exist. Thus, the following is true:

$$AXB + CXD = E \tag{2.13}$$

$$AXB + eAXD - eAXD + CXD = E \tag{2.13a}$$

$$AX(B+eD) + (C-eA)XD = E$$

$$(C-eA)^{-1}AX(B+eD) + XD = (C-eA)^{-1}E$$

$$(C-eA)^{-1}AX + XD(B+eD)^{-1} = (C-eA)^{-1}E(B+eD)^{-1}$$

Eq (2.13) has then been reduced to a Liapanov equation. By the results of Roth (Ref 38:392), the matrices in Eq (2.12) are similar which completes the proof.

Theorem 2.4: If $A+\lambda C$ and $B+\lambda D$ are regular pencils, then Eq (2.1) has a solution if and only if the following pair of matrices

$$\begin{bmatrix} (A+eC)^{-1}C & (A+eC)^{-1}E(D-eB)^{-1} \\ 0 & -D(B+eD)^{-1} \end{bmatrix}, \begin{bmatrix} (A+eC)^{-1} & 0 \\ 0 & -D(B+eD)^{-1} \end{bmatrix}$$

are similar, where there exists an e, a complex scalar, such that $|A+eC| \neq 0$ and $|D-eB| \neq 0$.

Proof: The proof of Theorem 2.4 is similar to the proof of Theorem 2.3. The major difference is that in Eq (2.13a), the matrix eCXB should be used instead of the matrix eAXD.

Theorem 2.5 (Sufficient Condition): Let

$$AX + CXB^{-1}D = E \tag{2.14}$$

where B is a nonsingular matrix, and let

$$\begin{bmatrix} A+\lambda C & E \\ 0 & -D+\lambda B \end{bmatrix}, \begin{bmatrix} A+\lambda C & 0 \\ 0 & -D+\lambda B \end{bmatrix} \tag{2.15}$$

be a pair of equivalent matrices, where $A+\lambda C$ and $D+\lambda B$ are regular pencils of matrices. Then there exists a solution of Eq (2.14).

Proof: By the results of Roth (Ref 38:392), and since the matrices in Eq (2.15) are equivalent with elements in $F(\lambda)$ where F is a polynomial domain and λ is an indeterminant, then $X(\lambda)$ and $Y(\lambda)$ are matrices with elements belonging to $F(\lambda)$. Hence Eq (2.14) can be rewritten as

$$(A+\lambda C)X(\lambda) - Y(\lambda)(D-\lambda B) = E \tag{2.16}$$

Let $$X(\lambda) = X_0+\lambda X_1+\lambda^2 X_2+\lambda^3 X_3+\ldots+\lambda^p X_p,\ p \leq m-1 \qquad (2.17)$$

and let

$$Y(\lambda) = Y_0+\lambda Y_1+\lambda^2 Y_2+\lambda^3 Y_3+\ldots+\lambda^q Y_q,\ q \leq n-1 \qquad (2.18)$$

where p and q denote the number of matrices with elements in F. Since $X(\lambda)$ and $Y(\lambda)$ are of the same dimension, then $p = q$. Thus Eq (2.16) can be expressed as

$$(A+\lambda C)(X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^p X_p) - (Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots+\lambda^p Y_p)(D-\lambda B)=E \qquad (2.19)$$

Equating the coefficients of like powers of λ in Eq (2.19), the following p+2 equations can be arrived at:

$$\begin{aligned}
AX_0 \quad &- \quad Y_0D \quad = E \\
AX_1 + CX_0 &- Y_1D + Y_0B = 0 \\
AX_2 + CX_1 &- Y_2D + Y_1B = 0 \\
&\vdots \qquad\qquad \vdots \qquad\quad \vdots \\
AX_h + CX_{h-1} &- Y_hD + Y_{h-1}B = 0 \\
&\vdots \qquad\qquad \vdots \qquad\quad \vdots \\
CX_p \quad &+ \quad Y_pB \quad = 0
\end{aligned} \qquad (2.20)$$

Multiplying each of the equations of Eq (2.20) by I, $B^{-1}D$, $(B^{-1}D)^2$, $(B^{-1}D)^3$,..., $(B^{-1}D)^{p+1}$ respectively yields:

$$\begin{aligned}
AX_0 \qquad\qquad\qquad\qquad &= E \\
AX_1(B^{-1}D) + CX_0(B^{-1}D) &= 0 \\
AX_2(B^{-1}D)^2 + CX_1(B^{-1}D)^2 &= 0 \\
\vdots
\end{aligned} \qquad (2.21)$$

$$
\begin{aligned}
AX_h(B^{-1}D)^h + CX_{h-1}(B^{-1}D)^h &= 0 \\
&\vdots \\
CX_p(B^{-1}D)^{p+1} &= 0
\end{aligned}
\tag{2.21}
$$

Adding the equations of Eq (2.21) and factoring yields:

$$
\begin{aligned}
&A[X_0+X_1(B^{-1}D)+X_2(B^{-1}D)^2+\ldots+X_p(B^{-1}D)^p] + \\
&\quad C[X_0(B^{-1}D)+X_1(B^{-1}D)^2+X_2(B^{-1}D)^3+\ldots+X_p(B^{-1}D)^{p+1}] = E
\end{aligned}
\tag{2.22}
$$

which implies

$$
\begin{aligned}
&A[X_0+X_1(B^{-1}D)+X_2(B^{-1}D)^2+\ldots+X_p(B^{-1}D)^p] + \\
&\quad C[X_0+X_1(B^{-1}D)+X_2(B^{-1}D)^2+\ldots+X_p(B^{-1}D)^p](B^{-1}D) = E
\end{aligned}
\tag{2.23}
$$

Hence, a solution of

$$AX + CXB^{-1}D = E \tag{2.24}$$

exists and the theorem is complete.

Conditions for the Solution of the Matrix Equation $\sum_{i=1}^{n} A_i X B_i = C$

Conditions that are necessary and sufficient for solutions of the general linear matrix equation

$$\sum_{i=1}^{n} A_i X B_i = C \tag{2.25}$$

will be based upon the ideas of a spectral decomposition of a matrix into a representation that uses idempotent and nilpotent matrices.

Definition: Let N be a square matrix with elements in a field F. If $N^2 = N$, N is said to be idempotent.

Definition: Let N be a square matrix with elements in a field F. If there exists a positive integer m such that $N^m = 0$, N is said to be nilpotent.

Definition: Let A and B be two idempotent matrices. Then if $AB = BA = 0$, A and B are said to be orthogonally idempotent.

Orthogonally idempotent matrices are generated by spectrally decomposing a matrix. Matrices that are associated with different eigenvalues are orthogonal.

Definition: Let $A = \sum_{i=1}^{n} \lambda_i E_i$. The representation of A where each λ_i is an eigenvalue of A and each E_i is an idempotent matrix associated with the λ_i is called a spectral decomposition.

From a theorem of Rosenblum (Ref 36:268), there exists the property that the sum of the orthogonal idempotent matrices equals the identity matrix, I.

As stated earlier, a considerable amount of work has been done to find the solution of equations of the type

$$BX - XA = Q \tag{2.26}$$

Most of this effort has been directed towards those cases in which the matrices A and B have been square and of the same dimension. Rosenblum (Ref 36) has derived both necessary and sufficient conditions for this case. The next theorem extends his work to the case where the matrices are of different

dimensions.

Theorem 2.6: Let A be a square matrix of dimension m, such that $A = \sum_{j=1}^{m} a_j E_j$ and B is a square matrix of dimension n, such that $B = \sum_{k=1}^{n} b_k F_k$, where the E_j's and F_k's each form distinct sets of orthogonal idempotent matrices. A necessary and sufficient condition that the matrix equation Eq (2.26) have a solution $X_{(n,m)}$ is that whenever for some pair of indices s and r, $a_s = b_r$, that is the characteristic roots are equal, then $F_r Q E_s = 0_{(n,m)}$.

Proof: To show necessity, suppose $X_{(n,m)}$ is a solution of Eq (2.26) such that $a_s = b_r$. Then Eq (2.26) implies

$$\sum_{k=1}^{n} (b_k F_k) X - X\left(\sum_{j=1}^{m} a_j E_j\right) = Q \qquad (2.27)$$

which can be rewritten as

$$\sum_{k=1}^{n} b_k F_k X - \sum_{j=1}^{m} a_j X E_j = 0 \qquad (2.28)$$

Multiplying from the left in a termwise fashion by F_r and at the same time multiplying from the right in a termwise fashion by E_s yields:

$$b_r F_r X E_s - a_s F_r X E_s = F_n Q E_s \qquad (2.29)$$

since the E_s's and F_r's are orthogonally idempotent. Factoring Eq (2.29) yields:

$$(b_r - a_s) F_r X E_s = F_r Q E_s \qquad (2.30)$$

But by hypothesis $b_r = a_s$, thus the left side of Eq (2.30)

becomes $0_{(n,m)}$. This then implies

$$F_r Q E_s = 0_{(n,m)}$$

To show sufficiency, the following convention will be used. If $a_s = b_r$, then $F_r Q E_s = 0$ and $\infty \cdot 0 = 0$. Thus the expression

$$\frac{1}{b_k - a_j} F_k Q E_j \tag{2.31}$$

has meaning for all j and k. Next let V be defined as follows:

$$V = \sum_{k=1}^{m} \sum_{j=1}^{n} \frac{1}{b_k - a_j} F_k Q E_j \tag{2.32}$$

If V is a solution of Eq (2.26), replacing X by V should result in Q. Hence,

$$\begin{aligned}
BV-VA &= B \sum_{k=1}^{m} \sum_{j=1}^{n} \frac{1}{b_k - a_j} F_k Q E_j - \sum_{k=1}^{m} \sum_{j=1}^{n} \frac{1}{b_k - a_j} F_k Q E_j A \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} \left(\frac{B F_k Q E_j - F_k Q E_j A}{b_k - a_j} \right) \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} \frac{\sum_{\lambda=1}^{} (b_\lambda F_\lambda) F_k Q E_j - F_k Q E_j \sum_{\mu=1}^{m} (a_\mu E_\mu)}{b_k - a_j} \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} \frac{b_k F_k Q E_j - a_j F_k Q E_j}{b_k - a_j} \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} F_k Q E_j \\
&= \left(\sum_{k=1}^{m} F_k \right) Q \left(\sum_{j=1}^{n} E_j \right) \\
&= Q
\end{aligned}$$

Therefore V is a solution of Eq (2.26) and the proof is complete.

Example 2.1: The technique derived in Theorem 2.6 will now be used to solve for the matrix X in the following equation:

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} X - X \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & -2 \end{bmatrix} \tag{2.33}$$

The representation of A is

$$A = a_1E_1 + a_2E_2 + a_3E_3$$

and B is

$$B = b_1F_1 + b_2F_2$$

In terms of the values from Eq (2.33), these become

$$A = 1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = 2 \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{2.34}$$

From Theorem 2.6 the solution will look like Eq (2.32), thus

$$X = \sum_{k=1}^{2} \sum_{j=1}^{3} \frac{F_kQE_j}{b_k - a_j}$$

$$X = \frac{F_1QE_1}{b_1-a_1} + \frac{F_1QE_3}{b_1-a_3} + \frac{F_2QE_1}{b_2-a_1} + \frac{F_2QE_2}{b_2-a_2}$$

$$X = \underbrace{\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 4 & 2 & -2\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}}_{2\text{-}1} + \underbrace{\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 4 & 2 & -2\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 1\end{bmatrix}}_{2\text{-}3} +$$

$$\underbrace{\begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 4 & 2 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}}_{3\text{-}1} + \underbrace{\begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 4 & 2 & -1\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 0\end{bmatrix}}_{3\text{-}2}$$

$$X = \begin{bmatrix}-3 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} - \begin{bmatrix}0 & 0 & 5\\ 0 & 0 & 0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}4 & 0 & 0\\ 4 & 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 2 & -2\\ 0 & 2 & -2\end{bmatrix}$$

$$X = \begin{bmatrix}-1 & 2 & -7\\ 2 & 2 & -2\end{bmatrix}$$

To extend the results of Theorem 2.6 use will be made of the fact that if two matrices commute, then the associated idempotent representation will also commute. The extended theorem is as follows.

Theorem 2.7: Let A and C be square matrices of dimension m, such that A and C can each be expressed as a sum of products of eigenvalues and orthogonal idempotent matrices. Let B and D be square matrices of dimension n, such that B and D can each be expressed as a sum of product of eigenvalue and orthogonal idempotent matrices. Also, let AC = CA and BD = DB. Then a necessary and sufficient condition that the matrix equation

$$BXC + DXA = Q \tag{2.35}$$

has a solution $X_{(n,m)}$ is that whenever for some set of indices {r,s,v,p} the following holds: $c_s b_v + a_r d_p = 0$, that the products of the characteristic roots sum to zero, then

$$H_p F_v Q G_s E_r = 0_{(n,m)} \tag{2.36}$$

Proof: To show necessity, suppose $X_{(n,m)}$ is a solution of Eq (2.35) such that $c_s b_v + a_r d_p = 0$, then Eq (2.35) implies

$$(\sum_{k=1}^{n} b_k F_k) X (\sum_{j=1}^{m} c_j G_j) + (\sum_{l=1}^{n} d_l H_l) X (\sum_{i=1}^{m} a_i E_i) = Q$$

Multiplying from the left by F_v in a termwise fashion and also multiplying from the right by G_s in a termwise fashion yields:

$$c_s b_v F_v X G_s + F_v (\sum_{l=1}^{n} d_l H_l) X (\sum_{i=1}^{m} a_i E_i) G_s = F_v Q G_s \tag{2.37}$$

Using the fact that if two matrices commute, then their idempotent matrices commute, Eq (2.37) can be written as

$$c_s b_v F_v X G_s + \sum_{l=1}^{n} d_l H_l \; F_v X G_s \sum_{i=1}^{m} a_i E_i = F_v Q G_s$$

Now multiply from the left by H_p in a termwise fashion and multiply from the right by E_r in a termwise fashion which results in

$$c_s b_v H_p F_v X G_s E_r + a_r d_p H_p F_v X G_s E_r = H_p F_v Q G_s E_r \tag{2.38}$$

Eq (2.38) then implies

$$(c_s b_v + a_r d_p) H_p F_v X G_s E_r = H_p F_v Q G_s E_r \tag{2.39}$$

By hypothesis $c_s b_v + a_r d_p = 0_{(n,m)}$ and thus the left hand side of Eq (2.39) is equal to $0_{(n,m)}$, which implies that

$$H_p F_v Q G_s E_r = 0_{(n,m)}$$

To show sufficiency, use will be made of the convention that whenever $c_s b_v = -a_r d_p$, then $H_p F_v X G_s E_r = 0$ and $\infty \cdot 0 = 0$, thus the expression

$$\frac{1}{c_j b_k + a_i d_l} H_l F_k Q G_j E_i$$

will have meaning. Define

$$V = \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{H_l F_k Q G_j E_i}{c_j b_k + a_i d_l} \tag{2.40}$$

and let V be a solution of Eq (2.35). Thus substituting the spectral representation for the matrices A,B,C, and D yields:

$$\begin{aligned} BVC+DVA = & \sum_{\lambda=1}^{n} b_\lambda F_\lambda \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{H_l F_k Q G_j E_i}{c_j b_k + a_i d_l} \sum_{\mu=1}^{m} c_\mu G_\mu + \\ & \sum_{\sigma=1}^{n} d_\sigma H_\sigma \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{H_l F_k Q G_j E_i}{c_j b_k + a_i d_l} \sum_{\phi=1}^{m} a_\phi E_\phi = \\ & \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{(\sum_{\lambda=1}^{n} b_\lambda F_\lambda) H_l F_k Q G_j E_i (\sum_{\mu=1}^{m} c_\mu G_\mu)}{c_j b_k + a_i d_l} + \\ & \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{(\sum_{\sigma=1}^{n} d_\sigma H_\sigma) H_l F_k Q G_j E_i (\sum_{\phi=1}^{m} a_\phi E_\phi)}{c_j b_k + a_i d_l} \end{aligned} \tag{2.41}$$

Using the properties of orthogonal idempotent matrices and commutativity, Eq (2.41) simplifies to

$$= \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{c_j b_k H_l F_k Q G_j E_i}{c_j b_k + a_i d_l} +$$

$$\sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{a_i d_l H_l F_k Q G_j E_i}{c_j b_k + a_i d_l}$$

$$= \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} \frac{(c_j b_k + a_i d_l) H_l F_k Q G_j E_i}{c_j b_k + a_i d_l}$$

$$= \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j=1}^{m} \sum_{l=1}^{n} H_l F_k Q G_j E_i$$

$$= \sum_{l=1}^{m} H_l \sum_{k=1}^{n} F_k \, Q \sum_{j=1}^{m} G_j \sum_{i=1}^{n} E_i$$

$$= Q$$

Hence V is a solution of Eq (2.35) and the theorem is proved.

In the more general case, solutions of equations of the type

$$\sum_{i=1}^{n} A_i X B_i = Q \qquad (2.42)$$

where the matrices A and B are square and of different dimension can be found. To solve this type of equation the following restrictions would have to hold:

i) $\{A_i\}$ would have to be a commutative set,

ii) $\{B_i\}$ would have to be a commutative set, and

iii) $\sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j = 0$

If these three restrictions hold, then the implication is that

$$\left(\sum_{i=1}^{n} E_i\right) Q \left(\sum_{j=1}^{m} F_j\right) = 0$$

Example 2.2: As an example of Theorem 2.7, consider the equation

$$BXA + BXA = Q \qquad (2.43)$$

where the B and A matrices are as defined in Example 2.1. Let

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

Then the solution for X would be:

$$X = \sum_{i=1}^{3} \sum_{k=1}^{2} \sum_{j=1}^{3} \sum_{l=1}^{2} \frac{H_l F_k Q G_j E_i}{c_j b_k + a_i d_l}$$

The summation would then involve summing 32 terms. This process could easily be adapted for a computer solution.

Not all matrices may be decomposed into a representation that involves only idempotent matrices. However, a matrix of this type may still be decomposed by using both idempotent matrices and nilpotent matrices. If E is an idempotent matrix, then let $\bar{E}$ be its associated nilpotent matrix. Two properties of nilpotent matrices are:

$$\text{i)} \quad E_i \bar{E}_i = \bar{E}_i = \bar{E}_i E_i$$

$$\text{ii)} \quad E_i \bar{E}_j = \bar{E}_i \bar{E}_j = 0 \quad (i \neq j) \qquad (2.44)$$

Generalizing Theorem 2.6 to this more inclusive case results in Theorem 2.8.

Theorem 2.8: Let A be a square matrix of dimension m, such that $A = \sum_{i=1}^{\bar{m}<m} a_i E_i + \bar{E}_i$ and let B be a square matrix of dimension n, such that $B = \sum_{j=1}^{\bar{n}<n} b_j F_j + \bar{F}_j$. The sets $\{E_i\}$, $\{F_j\}$ are complete sets of principal idempotent matrices and the sets $\{\bar{E}_i\}$, $\{\bar{F}_j\}$ are complete sets of nilpotent matrices associated with A and B respectively. If $Q\bar{E}_i = \bar{F}_j Q$ and

$X\bar{E}_i = \bar{F}_j X$ for all i and j, then a necessary and sufficient condition that $BX - XA = Q$ has a solution $X_{(n,m)}$ is that for some pair of integers r and s, $a_s = b_r$, then $B_r Q A_s = 0_{(n,m)}$.

Proof: To show necessity, let $X_{(n,m)}$ be a solution of

$$BX - XA = Q \tag{2.26}$$

such that $a_s = b_r$. then Eq (2.26) implies

$$[\sum_{j=1}^{\bar{n}\leq n} (b_j F_j + \bar{F}_j)]\, X - X\, [\sum_{i=1}^{\bar{m}\leq m} (a_i E_i + \bar{E}_i)] = Q$$

$$\sum_{j=1}^{\bar{n}\leq n} b_j F_j X + \sum_{j=1}^{\bar{n}\leq n} \bar{F}_j X - \sum_{i=1}^{\bar{m}\leq m} a_i X E_i - \sum_{i=1}^{\bar{m}\leq m} X\bar{E}_i = Q$$

Recalling the properties of nilpotent matrices that are stated in Eq (2.44) and multiplying from the left by F_r in a termwise fashion yields

$$b_r F_r X + \bar{F}_r X - \sum_{i=1}^{\bar{m}\leq m} a_i F_r X E_i - \sum_{i=1}^{\bar{m}\leq m} F_r X\bar{E}_i = F_r Q$$

Multiplying from the right by E_s in a termwise fashion yields

$$b_r F_r X E_s + \bar{F}_r X E_s - a_s F_r X E_s - F_r X\bar{E}_s = F_r Q E_s \tag{2.45}$$

Regrouping and simplifying Eq (2.45) becomes

$$(b_r - a_s) F_r X E_s + F_r(\bar{F}_r X - X\bar{E}_s) E_s = F_r Q E_s \tag{2.46}$$

But by hypothesis $\bar{F}_r X = X\bar{E}_s$ which implies Eq (2.46) is

$$(b_r - a_s) F_r X E_s = F_r Q E_s \tag{2.47}$$

Again, by hypothesis $b_r = a_s$ and hence the left side of

Eq (2.47) is $0_{(n,m)}$ which implies that

$$F_r Q E_s = 0_{(m,n)}$$

To show sufficiency the convention will be used that whenever $a_s = b_r$, then $F_r Q E_s = 0$ and $\infty \cdot 0 = 0$. Thus the expression

$$\frac{F_k Q E_j}{b_k - a_j}$$

will have meaning for all values of j and k.

Let V be a solution of Eq (2.26) where

$$V = \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{1}{b_k - a_i} F_k Q E_i \tag{2.48}$$

Then the left side of Eq (2.26) is

$$BV - VA = B\left(\sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{F_k Q E_i}{b_k - a_i}\right) - \left(\sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{F_k Q E_i}{b_k - a_i}\right) A$$

$$= \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{B F_k Q E_i}{b_k - a_i} - \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{F_k Q E_i A}{b_k - a_i} \tag{2.49}$$

In Eq (2.49) substitute the spectral decompositions for A and B to get

$$= \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{\sum_{\lambda=1}^{\bar{m}<m} (b_\lambda F_\lambda + \bar{F}_\lambda) F_K Q E_i}{b_k - a_i} - \sum_{k=1}^{\bar{m}<m} \sum_{i=1}^{\bar{n}<n} \frac{\sum_{\mu=1}^{\bar{n}<n} (a_\mu E_\mu + \bar{E}_\mu) F_k Q E_i}{b_k - a_i}$$

$$= \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{(b_k F_k + \bar{F}_k) Q E_i}{b_k - a_i} - \sum_{i=1}^{\bar{m}<m} \sum_{k=1}^{\bar{n}<n} \frac{F_k Q (a_i E_i + \bar{E}_i)}{b_k - a_i}$$

Adding and then factoring out the common terms yields:

$$= \sum_{i=1}^{\bar{m}\le m} \sum_{k=1}^{\bar{n}\le n} \frac{(b_k F_k + \bar{F}_k)QE - F_k Q(a_i E_i + \bar{E}_i)}{b_k - a_i} \qquad (2.50)$$

Expanding the numerator of Eq (2.50) changes the equation to

$$= \sum_{i=1}^{\bar{m}\le m} \sum_{k=1}^{\bar{n}\le n} \frac{b_k F_k QE_i + \bar{F}_k QE_i - a_i F_k QE_i - F_k Q\bar{E}_i}{b_k - a_i} \qquad (2.51)$$

Applying the hypothesis and the properties of idempotent and nilpotent matrices to Eq (2.51), this equation simplifies to:

$$= \sum_{i=1}^{\bar{m}\le m} \sum_{k=1}^{\bar{n}\le n} \frac{b_k F_k QE_i - a_i F_k QE_i}{b_k - a_i}$$

$$= \sum_{i=1}^{\bar{m}\le m} \sum_{k=1}^{\bar{n}\le n} \frac{(b_k - a_i) F_k QE_i}{b_k - a_i}$$

$$= \sum_{i=1}^{\bar{m}\le m} \sum_{k=1}^{\bar{n}\le n} F_k QE_i$$

$$= \left(\sum_{k=1}^{\bar{n}\le n} F_k \right) Q \left(\sum_{i=1}^{\bar{m}\le m} E_i \right)$$

$$= Q$$

Hence V is a solution of BX-XA=Q, and the theorem is complete.

Example: As an example of the procedures demonstrated in Theorem 2.8, let the matrices of Eq (2.26) be as follows:

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 3 & -2 \\ 4 & 1 & 0 \end{bmatrix} \text{ (Ref 4:186) and } B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \text{(Ref 29:277)} \qquad (2.52)$$

$$Q = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -1 \end{bmatrix}$$

Thus Eq (2.26) becomes

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} X - X \begin{bmatrix} 2 & -1 & 1 \\ 3 & 3 & -2 \\ 4 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -2 \\ & & \\ 2 & 1 & -1 \end{bmatrix} \tag{2.53}$$

The decompositions of A and B are:

$$A = a_1 E_1 + \bar{E}_1 + a_2 E_2 + \bar{E}_2$$

$$B = b_1 F_1 + b_2 F_2 \tag{2.54}$$

In terms of the specific A and B of Eq (2.52)

$$A = 1\left(\frac{1}{4}\begin{bmatrix} 2 & 2 & -2 \\ -1 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix}\right) + \left(\frac{1}{4}\begin{bmatrix} 0 & 0 & 0 \\ 10 & 10 & -10 \\ 10 & 10 & -10 \end{bmatrix}\right) + 3\left(\frac{1}{4}\begin{bmatrix} 2 & -2 & 2 \\ 1 & -1 & 1 \\ 3 & -3 & 3 \end{bmatrix}\right) + 0$$

$$B = -1\left(\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right) + 3\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) \tag{2.55}$$

From Theorem 2.8 the value of X that is being sought can be represented as Eq (2.48) as follows

$$X = \sum_{i=1}^{1<2} \sum_{k=1}^{0<2} \frac{F_k Q E_i}{b_k - a_i} \tag{2.56}$$

Expanding Eq (2.56) implies

$$X = \frac{F_1 Q E_1}{b_1 - a_1} + \frac{F_1 Q E_2}{b_1 - a_2} + \frac{F_2 Q E_1}{b_2 - a_1} \tag{2.57}$$

For all other choices of i and k the difference $b_k - a_i$ goes to zero and, hence, by the convention in Theorem 2.8 the quantity goes to zero. Making the substitutions into Eq (2.57)

$$X = (-\tfrac{1}{2})(\tfrac{1}{2})(\tfrac{1}{4})\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 2 & -2 \\ -1 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix}$$

$$+ (-\tfrac{1}{4})(\tfrac{1}{2})(\tfrac{1}{4})\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & -2 & 2 \\ 1 & -1 & 1 \\ 3 & -3 & 3 \end{bmatrix}$$

$$+ (\tfrac{1}{2})(\tfrac{1}{2})(\tfrac{1}{4})\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 2 & -2 \\ -1 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix} \qquad (2.58)$$

$$X = (-\tfrac{1}{16})\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + (-\tfrac{1}{32})\begin{bmatrix} -4 & 4 & -4 \\ 4 & -4 & 4 \end{bmatrix} + (\tfrac{1}{16})\begin{bmatrix} 12 & 12 & -12 \\ 12 & 12 & -12 \end{bmatrix}$$

$$X = \tfrac{1}{8}\begin{bmatrix} 7 & 5 & -5 \\ 5 & 7 & -7 \end{bmatrix}$$

The last theorem in this section generalizes the results of Theorem 2.7.

Theorem 2.9: Let A and C be square matrices of dimension m, such that $A = \sum_{i=1}^{\bar{m}<m} a_i E_i + \bar{E}_i$ and $C = \sum_{j=1}^{\bar{m}<m} c_j G_j + \bar{G}_j$, and let B and D be square matrices of dimention n, such that $B = \sum_{k=1}^{\bar{n}<n} b_k F_k + \bar{F}_k$ and $D = \sum_{l=1}^{\bar{n}<n} d_l H_l + \bar{H}_l$, where the sets $\{E_i\}$, $\{G_j\}$, $\{F_k\}$, and $\{H_l\}$ are complete collections of principal idempotent matrices and the sets $\{\bar{E}_l\}$, $\{\bar{G}_j\}$, $\{\bar{F}_k\}$, and $\{\bar{H}_l\}$ are complete collections of nilpotent matrices. Then a necessary and sufficient condition that the matrix equation

$$BXC + DXA = Q \qquad (2.35)$$

have a solution $X_{(n,m)}$ is that whenever for some set of indices $\{r,s,v,p\}$ where $H_v F_r$ and $G_s E_p$ form nonsingular matrices,

and

$$c_s\bar{F}_rQ + b_rQ\bar{G}_s + a_p\bar{H}_vQ + d_v\bar{E}_pQ + \bar{F}_rQ\bar{G}_s + \bar{H}_vQ\bar{E}_p = 0$$

implies that X is also a solution of

$$(b_rc_s+d_va_p+c_s\bar{F}_r+a_p\bar{H}_v)X + X(b_r\bar{G}_s+d_v\bar{E}_p) + \bar{F}_rX\bar{G}_s + \bar{H}_vX\bar{E}_p = Q$$

Proof: Let $X_{(n,m)}$ be a solution of Eq (2.35) such that

$$Q = \left(\sum_{k=1}^{\bar{n}<n} (b_kF_k+\bar{F}_k)\right)X\left(\sum_{j=1}^{\bar{m}<m} (c_jG_j+\bar{G}_j)\right)$$
$$+\left(\sum_{l=1}^{\bar{n}<n} (d_lH_l+\bar{H}_l)\right)X\left(\sum_{i=1}^{\bar{m}<m} (a_iE_i+\bar{E}_i)\right) \qquad (2.59)$$

Expanding Eq (2.59) yields

$$\begin{aligned} Q = &\left(\sum_{k=1}^{\bar{n}<n} b_kF_k\right)X\left(\sum_{j=1}^{\bar{m}<m} c_jG_j\right)+\left(\sum_{k=1}^{\bar{n}<n} b_kF_k\right)X\left(\sum_{j=1}^{\bar{m}<m} \bar{G}_j\right)\\ &+\left(\sum_{k=1}^{\bar{n}<n} \bar{F}_k\right)X\left(\sum_{j=1}^{\bar{m}<m} c_jG_j\right)+\left(\sum_{k=1}^{\bar{n}<n} \bar{F}_k\right)X\left(\sum_{j=1}^{\bar{m}<m} \bar{G}_j\right)\\ &+\left(\sum_{l=1}^{\bar{n}<n} d_lH_l\right)X\left(\sum_{i=1}^{\bar{m}<m} a_iE_i\right)+\left(\sum_{l=1}^{\bar{n}<n} d_lH_l\right)X\left(\sum_{i=1}^{\bar{m}<m} \bar{E}_l\right)\\ &+\left(\sum_{l=1}^{\bar{n}<n} \bar{H}_l\right)X\left(\sum_{i=1}^{\bar{m}<m} a_iE_i\right)+\left(\sum_{l=1}^{\bar{n}<n} \bar{H}_l\right)X\left(\sum_{i=1}^{\bar{m}<m} \bar{E}_i\right) \end{aligned} \qquad (2.60)$$

Multiplying on the left by F_r and H_v in that order in a termwise fashion and then multiplying on the right by G_s and E_p in that order in a termwise fashion making use of commutativity and the idempotent properties of these matrices yields from Eq (2.60)

$$H_vF_rQG_sE_p = b_rc_sH_vF_rXG_sE_p + b_rH_vF_rX\bar{G}_sE_p$$
$$+ C_sH_v\bar{F}_rXG_sE_p + H_v\bar{F}_rX\bar{G}_sE_p + d_va_pH_vF_rXG_sE_p \qquad (2.61)$$
$$+ d_vH_vF_rXG_s\bar{E}_p + a_p\bar{H}_vF_rXG_sE_p + \bar{H}_vF_rXG_s\bar{E}_p$$

Using the properties of nilpotent matrices and factoring, Eq (2.61) can be reduced to

$$H_vF_rQG_sE_p = H_vF_r[(b_rc_s+d_va_p+c_s\bar{F}_r+a_p\bar{H}_v)X$$
$$+ X(b_r\bar{G}_s+d_v\bar{E}_p) + \bar{F}_rX\bar{G}_s + \bar{H}_vX\bar{E}_p]G_sE_p \qquad (2.62)$$

Since H_vF_r and G_sE_p are nonsingular matrices, then their inverses exist and Eq (2.62) can be reduced to

$$Q = (b_rc_s+d_va_p+c_s\bar{F}_r+a_p\bar{H}_v)X+X(b_r\bar{G}_s+d_v\bar{E}_p)+\bar{F}_rX\bar{G}_s+\bar{H}_vX\bar{E}_p$$

Which implies X is also a solution of this equation and necessity has been shown. To show sufficiency the convention will be used that whenever $b_rc_s + d_va_p = 0$, then $H_vF_rXG_sE_p = 0$ and $\infty \cdot 0 = 0$, then the expression

$$\frac{H_vF_rXG_sE_p}{b_rc_s+d_va_p} \qquad (2.63)$$

will be well defined. Let V be a solution to Eq (2.63) where

$$V = \sum_{p=1}^{\bar{m}<m}\sum_{v=1}^{\bar{m}<m}\sum_{r=1}^{\bar{n}<n}\sum_{s=1}^{\bar{n}<n}\frac{H_vF_rQG_sE_p}{b_rc_s+d_va_p} \qquad (2.64)$$

Then

$$\text{BVC} + \text{DVA} = \left(\sum_{k=1}^{\bar{n}\le n}(b_kF_k+\bar{F}_k)\right)\left(\sum_{p=1}^{\bar{m}\le m}\sum_{v=1}^{\bar{m}\le m}\sum_{r=1}^{\bar{n}\le n}\sum_{s=1}^{\bar{n}\le n}\frac{H_vF_rQG_sE_p}{b_rc_s+d_va_p}\right)$$

$$\cdot\left(\sum_{j=1}^{\bar{m}\le m}(c_jG_j+\bar{G}_j)\right)+\left(\sum_{l=1}^{\bar{n}\le n}(d_lH_l+\bar{H}_l)\right)$$

$$\cdot\left(\sum_{p=1}^{\bar{m}\le m}\sum_{v=1}^{\bar{m}\le m}\sum_{r=1}^{\bar{n}\le n}\sum_{s=1}^{\bar{n}\le n}\frac{H_vF_rQG_sE_p}{b_rc_s+d_va_p}\right)\left(\sum_{i=1}^{\bar{m}\le m}(a_iE_i+\bar{E}_i)\right) \qquad (2.65)$$

Simplifying Eq (2.65) by multiplying and making use of the properties of idempotent and nilpotent matrices yields:

$$\text{BVC} + \text{DVA} = \sum_{p=1}^{\bar{m}\le m}\sum_{v=1}^{\bar{m}\le m}\sum_{r=1}^{\bar{n}\le n}\sum_{s-1}^{\bar{n}\le n}\left[\frac{(c_sb_r+d_va_p)\ H_vF_rQG_sE_p}{b_rc_s+d_va_p}\right.$$

$$\left.\frac{H_vF_r(c_sF_rQ+b_rQG_s+a_pH_vQ+d_lE_p+F_rQG_s+H_lQE_p)G_sE_p}{b_rc_s+d_va_p}\right] \qquad (2.66)$$

Utilizing the convention, Eq (2.66) simplifies to

$$= \sum_{p=1}^{\bar{m}\le m}\sum_{v=1}^{\bar{m}\le m}\sum_{r=1}^{\bar{n}\le n}\sum_{s=1}^{\bar{n}\le n}H_vF_rQG_sE_p$$

$$= Q$$

Which completes the proof of this theorem.

Generalizing these results to equations of the type

$$\sum_{i=1}^{n} A_iXB_i = Q \qquad (2.67)$$

where

$$A_i = \sum_{j=1}^{m}(a_{ij}A_{ij}+\bar{A}_{ij}) \text{ and } B_i = \sum_{j=1}^{m}(b_{ij}B_{ij}+\bar{B}_{ij})$$

with $\{A_{ij}\}$, $\{B_{ij}\}$ forming sets of complete idempotent matrices and $\{\bar{A}_{ij}\}$, $\{\bar{B}_{ij}\}$ forming complete sets of nilpotent matrices. A necessary and sufficient condition that the Eq (2.67) have a solution $X_{(n,m)}$ is that whenever for some set of indices the $\prod_{i=1}^{n}\prod_{j=1}^{m} A_{ij}$ and $\prod_{i=1}^{n}\prod_{j=1}^{m} B_{ij}$ have inverses and $\sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}\bar{A}_{ij}Q$

$+ \sum_{i=1}^{n}\sum_{j=1}^{m} \bar{A}_{ij}Q\bar{B}_{ij} = 0$, then $X_{(n,m)}$ is also a solution of

$$\left(\sum_{i=1}^{n}\prod_{j=1}^{m} a_{ij}b_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}\bar{A}_{ij}\right)X + X\left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\bar{B}_{ij}\right)$$

$$+ \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\prod_{i=1}^{n}\prod_{j=1}^{m} \bar{A}_{ij}X\bar{B}_{ij}\right) = Q.$$

An alternate solution to Eq (2.35) exists if either pair of matrices B,C or A,D are nonsingular. If this condition exists then the equation to be solved is

$$X + B^{-1}DXAC^{-1} = B^{-1}QC^{-1}$$

or

$$D^{-1}BXCA^{-1} + X = D^{-1}QA^{-1}$$

Spectral Decomposition of Rectangular Matrices

In this section, the results of the last section will be generalized to solutions of matrix equations in which the coefficients are rectangular matrices. In solving these equations use will be made of the following theorem.

Theorem 2.10: (Rao and Mitra (Ref 14:38)) Any m x n matrix A can be written as:

$$A = \sum_{i=1}^{m} \alpha_i U_i \tag{2.68}$$

where α_i^2, i=1, 2, ..., u are the distinct nonnull eigenvalues of A*A and the matrices

$$U_i = \alpha_i^{-1} A[I-(A^*A-\alpha_i^2 I)\{(A^*A-\alpha_i^2 I)^2\}^{-}(A^*A-\alpha_i^2 I)] \tag{2.69}$$

satisfy

$$U_i U_i^* U_i = U_i \; \forall i$$

$$U_i U_j^* = 0 \;, \; U^*U = 0 \; \forall i \neq j \tag{2.70}$$

(α_i is taken to be the positive square root of α_i^2 which is real and positive since A*A is hermitian and non-negative definite.)

For a proof of the above theorem, the reader is directed to the source cited. The notation used above is defined as follows: a matrix denoted by A* is the conjugate transpose of the matrix A; and a matrix A^- is the generalized inverse of the matrix A.

Example 2.3: As an example of Theorem 2.10 consider

$$A_{(2,3)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.71}$$

First the nonnull eigenvalues must be found, thus $A^* = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$ and the product

$$A^*A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \tag{2.72}$$

Therefore,

$$A^*A - \lambda I = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2-\lambda & 2 & 2 \\ 2 & 2-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{bmatrix} \tag{2.73}$$

$$= -\lambda^3 + 6\lambda^2$$

$$= \lambda^2(\lambda - 6)$$

Setting this last equation equal to zero and solving implies

$$\lambda = 0 \quad \text{or} \quad \lambda = 6$$

Since the α_i^2 are the nonnull values of λ, the only λ that can be set equal to α_i^2 is the value of 6. Thus $\alpha_i^2 = 6$ and hence

$$A = 6U_1 \tag{2.74}$$

To find the matrix U_1, Eq (2.69) must be solved. Making the substitution for α_i yields

$$U_1 = \frac{1}{\sqrt{6}} A[I - (A^*A - 6I)\{(A^*A - 6I)^2\}^- \quad (A^*A - 6I)] \tag{2.75}$$

$$A^*A - 6I = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \tag{2.76}$$

and

$$(A^*A - 6I)^2 = \begin{bmatrix} 24 & -12 & -12 \\ -12 & 24 & -12 \\ -12 & -12 & 24 \end{bmatrix} \tag{2.77}$$

with

$$\{(A^*A - 6I)^2\}^- = \begin{bmatrix} 1/6 & -1/12 & -1/12 \\ -1/12 & 1/6 & -1/12 \\ -1/12 & -1/12 & 1/6 \end{bmatrix} \tag{2.78}$$

Substituting Eqs (2.76),(2.77) and (2.78) into Eq (2.75)

$$U_1 = \frac{1}{\sqrt{6}}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

$$\cdot\left(\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} - \begin{bmatrix}-4 & 2 & 2\\ 2 & -4 & 2\\ 2 & 2 & -4\end{bmatrix}\begin{bmatrix}1/6 & -1/12 & -1/12\\ -1/12 & 1/6 & -1/12\\ -1/12 & -1/12 & 1/6\end{bmatrix}\begin{bmatrix}-4 & 2 & 2\\ 2 & -4 & 2\\ 2 & 2 & -4\end{bmatrix}\right)$$

$$= \frac{1}{\sqrt{6}}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\left(\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} - \begin{bmatrix}6 & -3 & -3\\ -3 & 6 & -3\\ -3 & -3 & 6\end{bmatrix}\right)$$

$$= \frac{1}{\sqrt{6}}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}-5 & 3 & 3\\ 3 & -5 & 3\\ 3 & 3 & -5\end{bmatrix}$$

$$= \frac{1}{\sqrt{6}}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

Thus A can be decomposed into a summation of products of scalers times matrices, where the scalers are determined by the eigenvalues of A*A.

Definition (Ref 35:20): Let A be an m x n matrix of arbitrary rank. A generalized inverse of A is an n x m matrix G such that $\vec{x} = G\vec{y}$ is a solution of $A\vec{x} = \vec{y}$ for any y which makes the equation consistent. One of the important properties of generalized inverses is found in the next lemma.

Lemma 2.1 (Rao and Mitra (Ref 35:20)): A^- exist if and only if $AA^-A = A$.

Example 2.4: Let A be given as in Eq (2.71). To find the generalized inverse of A the procedure outlined by Noble (Ref 30:339-341) will be followed. Let A be partitioned as follows:

$$A = \left[\begin{array}{c|cc}1 & 1 & 1\\ \hline 1 & 1 & 1\end{array}\right] \qquad (2.79)$$

then

$$A_{11} = [1] \ , \ A_{12} = [1 \ 1] \ , \ A_{21} = [1] \ , \text{ and } A_{22} = [1 \ \ 1] \quad (2.80)$$

From the matrices in Eq (2.80), the matrices Q, B, and C can be generated.

$$Q = A_{11}^{-1}A_{12} = [1][1 \ 1] = [1 \ 1]$$

$$B = \begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (2.81)$$

$$C = [I \ Q] = [1 \ 1 \ 1]$$

A formulation of the generalized inverse is then given by the equation

$$A^{-} = C^{T}(CC^{T})^{-1}(B^{T}B)^{-1}B^{T} \quad (2.82)$$

where the superscript T indicates the transpose matrix. Making the appropriate substitutions

$$A^{-} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \left([1 \ 1 \ 1] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)^{-1} \left([1 \ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} [1 \ 1]$$

$$A^{-} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} (1/3) \ (1/2) \ [1 \ 1]$$

$$A^{-} = \frac{1}{6} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.83)$$

Theorem 2.11 (Rao and Mitra (Ref 35:24): A necessary and sufficient condition for the equation AXB = C to have a solution is that

$$AA^{-}CB^{-}B = C \quad (2.84)$$

in which case the general solution is

$$X = A^{-}CB^{-}+Z-A^{-}AZBB^{-} \tag{2.85}$$

where Z is an arbitrary matrix.

Proof: Let there exist a matrix X such that AXB = C. Then:

$$\begin{aligned} AA^{-}CB^{-}B &= AA^{-}AXBB^{-}B \\ &= AXB \\ &= C \end{aligned}$$

Thus necessity follows. Sufficiency is obvious since $A^{-}CB^{-}$ is clearly a solution.

Example 2.5: Let AXB = C where $A_{(2,3)}$, $B_{(3,4)}$ and $C_{(2,4)}$ are defined as follows:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 2 & 3 & -1 & -4 \\ 3 & 4 & -1 & -5 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 3 & 2 & -1 & 0 \\ 4 & 0 & 1 & 1 \end{bmatrix}$$

then solve for the matrix X.

According to Theorem 2.11, all that is required to solve for X is to set X equal to the result of Eq (2.85). Doing this will generate a family of solutions dependent on the choice of the matrix Z. For purposes of this example, let Z be equal to the zero matrix, then $X = A^{-}CB^{-}$. A^{-} is as found in Example 2.4, Eq (2.83).

To find the generalized inverse of B, the same procedure will be followed. Checking the rank of B, it is easily shown that $\rho(B) = 2$, and thus the partition of B is

$$B = \left[\begin{array}{cc:cc} 1 & 1 & 0 & -1 \\ 2 & 3 & -1 & -4 \\ \hdashline 3 & 4 & -1 & -5 \end{array}\right] \tag{2.86}$$

Therefore

$$B_{11} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix},\ B_{12} = \begin{bmatrix} 0 & -1 \\ -1 & -4 \end{bmatrix},\ B_{21} = [3\ 4]\ ,\ \text{and}\ B_{22} = [-1\ -5]$$

The intermediate matrices Q, B', and C' are

$$Q = [B_{11}^{-1}][B_{12}] = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ -1 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$$

$$B' = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \tag{2.87}$$

and

$$C' = [I\ Q] = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -2 \end{bmatrix}$$

Thus

$$B^- = C'^T(C'C'^T)^{-1}(B'^TB')^{-1}B'^T$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & -2 \end{bmatrix}\left(\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & -2 \end{bmatrix}\right)^{-1}\left(\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

$$B^- = (1/9)\begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 7 & -5 & 2 \\ -5 & 4 & -1 \end{bmatrix}$$

$$B^- = (1/9)\begin{bmatrix} 9 & 6 & 3 \\ 2 & -1 & 1 \\ 7 & -5 & 2 \\ 5 & -4 & 1 \end{bmatrix} \tag{2.88}$$

Returning to the solution of X and making the appropriate substitutions from Eqs (2.83) and (2.88) yields

$$X = A^-CB^-$$

$$X = \left((1/6)\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\right)\begin{bmatrix}3 & 2 & -1 & 0\\ 4 & 0 & 1 & 1\end{bmatrix}\left((1/9)\begin{bmatrix}9 & 6 & 3\\ 2 & -1 & 1\\ 7 & -5 & 2\\ 5 & -4 & 1\end{bmatrix}\right)$$

which when simplified yields:

$$X = (1/9)\begin{bmatrix}12 & 6 & 4\\ 12 & 6 & 4\\ 12 & 6 & 4\end{bmatrix}$$

The next theorem generalizes the conclusions of the last theorem.

Theorem 2.12: A necessary and sufficient condition for the equation

$$AXC + DXB = Q \tag{2.89}$$

to have a solution is that $RR^*WL^*L = W$, where $W = RR^*QL^*L$ and R and L are spectrally decomposed matrices multiplied from the right and left respectively. Then a general solution of Eq (2.89) is

$$X = R^-WL^- + Z - R^-RZLL^-$$

where Z is an arbitrary matrix. This solution exists provided at least one of the following statements holds, but not both i and ii at the same time:

$$\begin{aligned} &\text{i)} \quad U_iU_i^*D = 0 \quad \text{or} \quad BW_k^*W_k = 0\\ &\text{ii)} \quad Y_1Y_1^*A = 0 \quad \text{or} \quad CV_j^*V_j = 0 \end{aligned} \tag{2.90}$$

where the decomposition of the matrices A, B, C, and D are as follows:

$$A_{(m,n)} = \sum_{i=1}^{e}\alpha_iU_i \ , \ B_{(p,q)} = \sum_{j=1}^{f}\beta_jV_j \ , \ C_{(p,q)} = \sum_{k=1}^{g}\gamma_kG_k$$

and

$$D_{(m,n)} = \sum_{l=1}^{h} \delta_l Y_l$$

Proof: If Eq (2.89) is a true statement, then

$$\left(\sum_{i=1}^{e} \alpha_i U_i\right) X \left(\sum_{k=1}^{g} \gamma_k G_k\right) + \left(\sum_{l=1}^{h} \delta_l Y_l\right) X \left(\sum_{j=1}^{f} \beta_j V_j\right) = Q$$

Multiplying on the left by $U_i U_i^*$ and from the right by $G_k G_k^*$ yields

$$U_i X G_k + (U_i U_i^* \sum_{i=1}^{h} \delta_l Y_l) X (\sum_{j=1}^{f} \beta_j V_j G_k^* G_k) = U_i U_i^* Q G_k^* G_k$$

which can be simplified to

$$U_i X G_k + U_i U_i D X B G_k^* G_k = U_i U_i^* Q G_k^* G_k$$

But by hypothesis (i) either $U_i U_i^* D = 0$ or/and $B G_k^* G_k = 0$, then the last equation reduces to

$$U_i X G_k = U_i U_i^* Q G_k^* G_k$$

Then this is a form of Theorem 2.11 and the proof of the theorem is complete.

Note that if the multiplication had been by $Y_1^* Y_1$ and $V_j V_j^*$ respectively, then the results would have been

$$\delta_1 \beta_j Y_1 X V_j = Y_1 Y_1^* Q V_j^* V_j \qquad (2.91)$$

and either $Y_1 Y_1^* A = 0$ or $C V_j^* V_j = 0$ would have to be true for the theorem to hold. Thus there exists four distinct ways for this theorem to hold.

Example 2.6: Consider the equation,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} X \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} X \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \quad (2.92)$$

which corresponds to Eq (2.89). Then $A^* = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$ and from Eq (2.88) $A^- = (1/6)\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$

Decomposing A yields

$$A = \alpha_i U_i = (\sqrt{6})(1/\sqrt{6})\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.93)$$

which implies

$$\alpha_i = \sqrt{6} \text{ and } U_i = (1/\sqrt{6})\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

and

$$UU^* = (1/2)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.94)$$

Following the same procedure for the matrix D implies

$$D^* = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \text{ and } D^- = (1/4)\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Decomposing D yields

$$D = \delta_k Y_k = (2)(1/2)\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \quad (2.95)$$

which implies

$$\delta_k = 2 \quad \text{and} \quad Y_k = (1/2)\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

thus

$$Y^*Y = (1/2)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.96)$$

Substituting the decompositions of A and D from Eqs (2.93) and (2.95) respectively into Eq (2.92) yields

$$(\sqrt{6})(1/\sqrt{6})\begin{bmatrix}1 & 1 & 1\\ & & \\ 1 & 1 & 1\end{bmatrix}X\ (2)(1/2)\begin{bmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{bmatrix}+\begin{bmatrix}1 & 0 & -1\\ & & \\ -1 & 0 & 1\end{bmatrix}X\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}=\begin{bmatrix}1 & 3\\ & \\ 2 & -1\end{bmatrix}$$

Multiplying on the left by UU* from Eq (2.94) and on the right by Y*Y from Eq (2.96) yields

$$(1/2)\begin{bmatrix}1 & 1\\ & \\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\ & & \\ 1 & 1 & 1\end{bmatrix}X\ (1/2)\begin{bmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ & \\ 1 & 1\end{bmatrix}$$

$$+\ (1/2)\begin{bmatrix}1 & 1\\ & \\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & -1\\ & & \\ -1 & 0 & 1\end{bmatrix}X\ (1/2)\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ & \\ 1 & 1\end{bmatrix}$$

$$=\ (1/2)(1/2)\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 3\\ 2 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} \qquad (2.97)$$

In simplifying this last equation (2.97), the second term on the left goes to zero, thus allowing the use of Theorem 2.12. The simplified form of Eq (2.97) is then

$$\begin{bmatrix}1 & 1 & 1\\ & & \\ 1 & 1 & 1\end{bmatrix}X\begin{bmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{bmatrix}+\ 0\ =\ (5/4)\begin{bmatrix}1 & 1\\ & \\ 1 & 1\end{bmatrix} \qquad (2.98)$$

Utilizing Theorem 2.12 the solution is

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

$$+\ Z-\begin{bmatrix}1 & 1 & 1\\ & & \\ 1 & 1 & 1\end{bmatrix}^{-}\begin{bmatrix}1 & 1 & 1\\ & & \\ 1 & 1 & 1\end{bmatrix}Z\begin{bmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{bmatrix}^{-} \qquad (2.99)$$

In Eq (2.99), letting Z = 0 implies

$$X=(5/24)\begin{bmatrix}1 & 0 & 1\\ 1 & 0 & 1\\ 1 & 0 & 1\end{bmatrix}$$

and if $Z=\begin{bmatrix}1 & 2 & 1\\ 3 & 4 & -2\\ -1 & -3 & 1\end{bmatrix}$, then the X is Eq (2.99) becomes

$$X = (1/24)\begin{bmatrix} 17 & 48 & 17 \\ 65 & 96 & -55 \\ -31 & -79 & 17 \end{bmatrix}$$

Solution of the Matrix Equation AXB + CYD = E

In this section equations of the type

$$AXB + CYD = E$$

will be considered. The theorems to follow will be proved for cases in which the variable matrices are the same. After the proof is complete, those changes that are necessary for a proof of the two variable case will be given.

Returning to the work of Roth (Ref 38), it should be noted that his results are all stated for cases in which the variable matrices are square. In the next four theorems these results will be extended to the more general case in which the variable matrices are rectangular.

Theorem 2.13: A necessary and sufficient condition that the matrix equation

$$AXI_1 - I_2XB = C \tag{2.100}$$

where I_1 and I_2 are identity matrices and A and I_2 are square matrices of dimension m and B and I_1 are square matrices of dimension n, all with elements in the field F, is that the matrices

$$\begin{bmatrix} A & C \\ O & B \end{bmatrix} \text{ and } \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.101}$$

be similar.

Proof: To show similarity, there must exist matrices P

$$P\begin{bmatrix} A & C \\ O & B \end{bmatrix}P^{-1} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.102}$$

Let $P = \begin{bmatrix} I_{(m,m)} & IX_{(m,n)} \\ O_{(n,m)} & I_{(n,n)} \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} I_{(m,m)} & -XI_{(m,n)} \\ O_{(n,m)} & I_{(n,n)} \end{bmatrix}$

Substituting the choice for P and P^{-1} into Eq (2.102) yields

$$\begin{bmatrix} I_{(m,m)} & IX_{(m,n)} \\ O_{(n,m)} & I_{(n,n)} \end{bmatrix}\begin{bmatrix} A_{(m,m)} & C_{(m,n)} \\ O_{(n,m)} & B_{(n,n)} \end{bmatrix}\begin{bmatrix} I_{(m,m)} & -XI_{(m,n)} \\ O_{(n,m)} & I_{(n,n)} \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

$$\begin{bmatrix} IAI & -IAXI+ICI+IXBI \\ O & IBI \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.103}$$

But from Eq (2.100) -IAXI+ICI+IXBI is equal to O. Therefore, Eq (2.103) can be rewritten as

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

To show sufficiency, reliance is made on the fact that since the matrices in Eq (2.101) are similar, then the following pair of matrices

$$\begin{bmatrix} A-\lambda I & C \\ O & B-\lambda I \end{bmatrix} \text{ and } \begin{bmatrix} A-\lambda I & O \\ O & B-\lambda I \end{bmatrix} \tag{2.104}$$

whose elements are in F[x] will also be similar. From Roth's Lemma (Ref 38:392) there exists matrices X and Y such that

$$X = X_0 + \lambda X_1 + \lambda^2 X_2 + \ldots + \lambda^n X_n$$

$$Y = Y_0 + \lambda Y_1 + \lambda^2 Y_2 + \ldots + \lambda^n X_n$$

and

$$(A - \lambda I_2)\, X - Y\,(B - \lambda I_1) = C \tag{2.105}$$

Substituting the values for X and Y into Eq (2.105) yields

$$(A-\lambda I_2)(X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^n X_n)I_1$$
$$-I_2\,(Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots+\lambda^n Y_n)(B-\lambda I_1) = C \quad (2.106)$$

After multiplying out Eq (2.106) and then equating like powers of λ, the following set of $\lambda+2$ equations is generated.

$$\begin{aligned}
&AX_0I_1 && -I_2Y_0B && = C\\
&AX_1I_1 + I_2X_0I_1 && -I_2Y_1B + I_2Y_0I_1 && = 0\\
&AX_2I_1 + I_2X_1I_1 && -I_2Y_2B + I_2Y_1I_1 && = 0 \quad (2.107)\\
&&&\vdots\\
&AX_nI_1 + I_2X_{n-1}I_1 && -I_2Y_nB + I\,Y_{n-1}I_1 && = 0\\
&\qquad I_2X_nI_1 && \qquad + I_2Y_nI_1 && = 0
\end{aligned}$$

All of the I_1 and I_2 in Eq (2.107) are square and thus can be dropped from the equation without loss of generality. Now multiply each row of Eq (2.107) by I, B, B^2, B^3, ..., B^{n+1} respectively and sum the members of the resulting equations. After factoring the result is

$$A(X_0+X_1B+X_2B^2+\ldots+X_nB^n) - (X_0+X_1B+X_2B^2+\ldots+X_nB^n)B = C \quad (2.108)$$

Multiplying from the right and left by I_2 and I_1 respectively yields

$$A(X_0+X_1B+X_2B^2+\ldots+X_nB^n)I_1 - I_2(X_0+X_1B+X_2B^2+\ldots+X_nB^n)B = C$$

which therefore implies that a solution of Eq (2.100) exists

and is

$$X = X_0 + X_1 B + X_2 B^2 + X_3 B^3 + \ldots + X_n B^n$$

By an entirely similar argument it can be shown that

$$A(Y_0 + Y_1 B + Y_2 B^2 + \ldots + Y_n B^n) I_1 - I_2 (Y_0 + Y_1 B + Y_2 B^2 + \ldots + Y_n B^n) B = C$$

also implies that a solution of Eq (2.100) exists and is

$$X = Y_0 + Y_1 B + Y_2 B^2 + Y_3 B^3 + \ldots + Y_n B^n$$

If Eq (2.100) had been

$$AXI_1 - I_2 YB = C$$

then the proof of the theorem would have been identical except that the matrices in Eq (2.101) would be equivalent, with matrices P and Q being defined as

$$P = \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \text{ and } Q = \begin{bmatrix} I & -XI \\ 0 & I \end{bmatrix}$$

The next theorem extends the last result by changing one of the identity matrices into a matrix that is not an identity.

Theorem 2.14: A necessary and sufficient condition that the equation

$$AX - DXB = C \qquad (2.109)$$

where A and D are m x m matrices and B is an n x n matrix with elements in F, have a solution X with elements in F is that the pair of matrices

$$\begin{bmatrix} A & C \\ O & B \end{bmatrix}, \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.110}$$

be equivalent.

Proof. Since the matrices of Eq (2.110) are to be equivalent, then there must exist matrices P and Q such that

$$P \begin{bmatrix} A & C \\ O & B \end{bmatrix} Q = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.111}$$

Therefore let

$$P = \begin{bmatrix} I & -DX \\ O & I \end{bmatrix} \text{ and } Q = \begin{bmatrix} I & X \\ O & I \end{bmatrix}$$

which transforms Eq (2.111) into

$$\begin{bmatrix} I & -DX \\ O & I \end{bmatrix} \begin{bmatrix} A & C \\ O & B \end{bmatrix} \begin{bmatrix} I & X \\ O & I \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.112}$$

Upon simplification, this becomes

$$\begin{bmatrix} IAI & IAX+ICI-DXBI \\ O & IB \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.113}$$

Since the identity matrices in Eq (2.113) are all square, no changes occur during multiplication, thus, for example, IAI=A Also making use of Eq (2.109), it should be noted that $AX + C - DXB = O$. Hence Eq (2.113) becomes

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

Since the necessary condition has been established, the matrix pair

$$\begin{bmatrix} A-\lambda D & C \\ O & B-\lambda I \end{bmatrix} \text{ and } \begin{bmatrix} A-\lambda D & O \\ O & B-\lambda I \end{bmatrix} \tag{2.114}$$

will also be equivalent with elements in F[x].

Following the same procedure as in Theorem 2.14, Roth's Lemma (Ref 38:392) states that there exists matrices X and Y such that

$$X = X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^n X_n$$

and

$$Y = Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots\lambda^n Y_n \tag{2.115}$$

Thus Eq (2.109) becomes

$$\begin{aligned}&(A-\lambda D)(X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^n X_n)\\ &- D(Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots\lambda^n Y_n)(B-\lambda I) = C\end{aligned} \tag{2.116}$$

Upon expansion and equating terms with like powers of λ the following system of $\lambda + 2$ equations is formed.

$$\begin{aligned}
AX_0 \qquad\qquad & -DY_0B \qquad\qquad && = C\\
AX_1 - DX_0 \quad & - DY_1B + Y_0I && = 0\\
AX_2 - DX_1 \quad & - DY_2B + Y_1I && = 0\\
& \vdots && \\
AX_n - DX_{n-1} \quad & - DY_nB + Y_{n-1}I && = 0\\
- DX_n \qquad\qquad & + Y_nI && = 0
\end{aligned} \tag{2.117}$$

Upon multiplying the system in Eq (2.117) by I, B, B^2, ..., B^{n+1} respectively, then adding columnwise and factoring out like terms yields

$$A(X_0+X_1B+X_2B^2+\ldots+X_nB^n)-D(X_0+X_1B+X_2B^2+\ldots+X_nB^n)B = C$$

This then implies that

$$X_0+X_1B+X_2B^2+X_3B^3+\ldots+X_nB^n$$

is a solution of Eq (2.109), which completes the proof.

By similar arguments it can be shown that necessary and sufficient conditions exist for a solution of the equation

$$AXE - XB = C \tag{2.118}$$

In showing these conditions exist, the matrix pair

$$\begin{bmatrix} A-\lambda I & C \\ O & B-\lambda E \end{bmatrix} \text{ and } \begin{bmatrix} A-\lambda I & O \\ O & B-\lambda E \end{bmatrix}$$

are seen to be equivalent and the system of equations to be solved is

$$\begin{aligned}
&AX_0E && - Y_0B && = C \\
&AX_1E - IX_0E && - Y_1B + Y_0E && = 0 \\
&AX_2E - IX_1E && - Y_2B + Y_1E && = 0 \\
&\qquad\vdots \\
&AX_nE - IX_{n-1}E && - Y_nB + Y_{n-1}E && = 0 \\
&\quad - IX_nE && + Y_nE && = 0
\end{aligned} \tag{2.119}$$

The $\lambda + 2$ equations of Eq (2.119) are then multiplied from the right by $I, A, A^2, \ldots, A^{n+1}$ respectively. The result is that $Y_0+AY_1A^2Y_2+A^3Y_3+\ldots+A^nY_n$ is also a solution of Eq (2.118).

If Eq (2.119) had been expressed as

$$AX - DYB = C$$

then the method of solution and the proof of Theorem 2.14 would still be valid. To show equivalence, the matrix P

would have to be changed to

$$P = \begin{bmatrix} I & -DY \\ O & I \end{bmatrix}$$

and the remainder of the proof follows in a similar fashion.

Theorem 2.15: A necessary and sufficient condition that

$$AXE - DXB = C \tag{2.120}$$

where A and D are square matrices of dimension m and B and E are square matrices of dimension n, all with elements in F, have a solution X with elements in F is that the matrix pair

$$\begin{bmatrix} A & C \\ O & B \end{bmatrix} \text{ and } \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.121}$$

be equivalent and that either

i) E be nonsingular and BE = EB

or ii) D be nonsingular and AD = DA

Proof: To show equivalence there must exist matrices P and Q such that

$$P \begin{bmatrix} A & C \\ O & B \end{bmatrix} Q \text{ and } \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.122}$$

Let

$$P = \begin{bmatrix} I & DX \\ O & I \end{bmatrix} \text{ and } Q = \begin{bmatrix} I & -XE \\ O & I \end{bmatrix}$$

then Eq (2.122) becomes

$$\begin{bmatrix} I & DX \\ O & I \end{bmatrix} \begin{bmatrix} A & C \\ O & B \end{bmatrix} \begin{bmatrix} I & -XE \\ O & I \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.123}$$

Simplifying and realizing that multiplication by an identity matrix leaves the original matrix unchanged yields

$$\begin{bmatrix} A - AXE + C + DXB & \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.124}$$

Using Eq (2.120), Eq (2.124) can be simplified to

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

To show sufficiency, reliance will be made upon what has just been proved. Since the matrices of Eq (2.121) are equivalent, then the matrix pair of Eq (2.125) will also be equivalent.

$$\begin{bmatrix} A-\lambda D & C \\ O & B-\lambda E \end{bmatrix} = \begin{bmatrix} A-\lambda D & E \\ O & B-\lambda E \end{bmatrix} \tag{2.125}$$

Utilizing Roth's Lemma (Ref 38:392), there exists matrices X and Y where X and Y are as expressed in Eq (2.115). Making appropriate substitutions in Eq (2.120) yields the equivalent equation

$$\begin{aligned} &(A-\lambda D)(X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^n X_n)E \\ &-D(Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots+\lambda^n Y_n)(B-\lambda E) = C \end{aligned} \tag{2.126}$$

Multiplying out Eq (2.126) and then equating coefficients of like powers of λ yields the system of equations

$$\begin{aligned} AX_0E \quad &- DY_0B \quad = C \\ AX_1E - DX_0E \; &- DY_1B + DY_0E = 0 \\ AX_2E - DX_1E \; &- DY_2B + DY_1E = 0 \\ &\vdots \end{aligned} \tag{2.127}$$

$$AX_nE - DX_{n-1}E - DY_nB + DY_{n-1}E = 0$$
$$- DX_nE \qquad\qquad + DY_nE \quad = 0$$

Next multiply each of the $\lambda + 2$ equations of Eq (2.127 by $I, E^{-1}B, (E^{-1}B)^2, (E^{-1}B)^3, \ldots (E^{-1}B)^{n+1}$ respectively and then sum columnwise to get as a result

$$A[X_0+X_1E^{-1}B+X_2(E^{-1}B)^2+\ldots+X_n(E^{-1}B)^n]E$$
$$-D[X_0+X_1E^{-1}B+X_2(E^{-1}B)^2+\ldots+X_n(E^{-1}B)^n]B = C$$

which implies that

$$X_0+X_1E^{-1}B+X_2(E^{-1}B)^2+\ldots+X_n(E^{-1}B)^n$$

is also a solution of $AXE - DXB = C$ which completes the proof.

Similar results hold when Eq (2.127) is multiplied by $I, AD^{-1}, (AD^{-1})^2, (AD^{-1})^3, \ldots, (AD^{-1})^n$ respectively to yield the equation

$$-D[Y_0+Y_1AD^{-1}+Y_2(AD^{-1})^2+\ldots+Y_n(AD^{-1})^n]B$$
$$+A(Y_0+Y_1AD^{-1}+Y_2(AD^{-1})^2+\ldots+Y_n(AD^{-1})^n]E = C \qquad (2.128)$$

Eq (2.128) can then be arranged to be in the form of Eq (2.120) which then implies that a second set of solutions can be generated by

$$Y_0+Y_1AD^{-1}+Y_2(AD^{-1})^2+\ldots+Y_n(AD^{-1})^n$$

The following example incorporates the technique used in the three previous theorems. To fit this example to any of the individual theorems, appropriate matrices would need to be deleted.

Stating Eq (2.120) as

$$AXE - DYB = C$$

only has minor effects upon the method of proof given in Theorem 2.15. The matrices can be shown to be equivalent if P is changed to

$$P = \begin{bmatrix} I & DY \\ 0 & I \end{bmatrix}$$

and the rest of the proof follows as discussed in the theorem.

Example 2.7: Consider the equation $AXI_1 - I_2YD = E$, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ I_1 = \begin{bmatrix} 1 & 0 \end{bmatrix},\ I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ D = \begin{bmatrix} 1 & 1 \end{bmatrix},\ \text{and } E = \begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix}$$

Thus the equation of the form expressed in Eq (2.126) is

$$(A-\lambda I)(X)I_1 - I_2(Y)(D-\lambda I) = E \qquad (2.129)$$

After making appropriate substitutions, the resulting equation is

$$\begin{bmatrix} 1-\lambda & 0 \\ 0 & -\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\begin{bmatrix} 1-\lambda & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & -1 \\ -1 & -1 \end{bmatrix}$$

which simplifies to

$$\begin{bmatrix} (1-\lambda)x_1 & 0 \\ -\lambda x_2 & 0 \end{bmatrix} - \begin{bmatrix} y_1(1-\lambda) & y_1 \\ y_2(1-\lambda) & y_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} \qquad (2.130)$$

Eq (2.130) gives rise to the solution of four equations in four unknowns. These equations are:

$$\begin{aligned} (1-\lambda)x_1 - y_1(1-\lambda) &= 0 \\ - y_1 &= -1 \\ -\lambda x_2 - y_2(1-\lambda) &= -1 \\ -y_2 &= -1 \end{aligned} \qquad (2.131)$$

This system of equations then quickly reduces to

$$x_1 = y_1 \quad , \quad y_1 = 1 \quad , \quad y_2 = 1$$

and upon further simplification, the result $x_2 = 1$ is achieved. Thus the solution to this equation is

$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The next theorem generalizes the last results to cases in which the coefficient matrices are non-square. In addition, the variable matrices will also be non-square. In this theorem the proof will be in terms of two variable matrices X and Y, but with minor changes the theorem is applicable to equations in which the variables are the same.

Theorem 2.16: Let $A_{(m,n)}$, $I_{2(m,n)}$, $B_{(p,q)}$, and $C_{(m,q)}$ be matrices with elements in the polynomial domain F[x] of a field and let I_1 and I_2 be identity matrices. A necessary and sufficient condition that the matrix equation

$$AXI_1 - I_2YB = C \qquad (2.132)$$

have a solution $X_{(n,p)}$, $Y_{(n,p)}$ with elements in F[x] is that the matrices

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \text{ and } \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

be equivalent.

Proof: To show equivalence there must exist matrices P and Q such that

$$P \begin{bmatrix} A & C \\ O & B \end{bmatrix} Q = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.133}$$

The choices of P and Q in Eq (2.133) are

$$P = \begin{bmatrix} I_{(m,m)} & I_{(m,n)}Y_{(n,p)} \\ O_{(p,m)} & I_{(p,p)} \end{bmatrix} \text{ and } Q = \begin{bmatrix} I_{(m,m)} & -X_{(n,p)}I_{(p,q,)} \\ O_{(q,n)} & I_{(q,q)} \end{bmatrix}$$

Thus Eq (2.133) becomes upon substitution

$$\begin{bmatrix} I & IY \\ O & I \end{bmatrix}\begin{bmatrix} A & C \\ O & B \end{bmatrix}\begin{bmatrix} I & -XI \\ O & I \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.134}$$

Performing the multiplications reduces Eq (2.134) to

$$\begin{bmatrix} A & -AXI+C+IYB \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.135}$$

Using Eq (2.127), Eq (2.130) can be reduced to

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

which proves necessity. Sufficiency will be proved with the aid of the idea of regular pencils. Thus Eq (2.132) can be written as

$$(A-\lambda I)XI_1 - I_2Y(B-\lambda I) = C \tag{2.136}$$

The matrices X and Y can also be written as

$$X = X_0+\lambda X_1+\lambda^2 X_2+\ldots+\lambda^p X_p$$

and

$$Y = Y_0+\lambda Y_1+\lambda^2 Y_2+\ldots+\lambda^p Y_p$$

Multiplying out Eq (2.137) and then equating like power of λ yields the following system of $\lambda + 2$ equations

$$\begin{aligned}
&AX_0I_1 && - I_2Y_0B && = C\\
&AX_1I_1 - IX_0I_1 && - I_2Y_1B + I_2Y_0I && = 0\\
&AX_2I_1 - IX_1I_1 && - I_2Y_2B + I_2Y_1I && = 0\\
&&&\vdots && \\
&AX_nI_1 - IX_{n-1}I_1 && - I_2Y_nB + I_2Y_{n-1}I && = 0\\
&\quad - IX_nI && && + I_2Y_nI = 0
\end{aligned} \qquad (2.138)$$

Multiply each equation of the system in Eq (2.138) by $I_{(q,p)}$, $I_{(q,p)}(BI)_{(p,p)}$, $I_{(q,p)}(BI)^2_{(p,p)}$,..., $I_{(q,p)}(BI)^{n+1}_{(p,p)}$ respectively, then sum the resulting equations and factor common terms to yield:

$$A[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_2(BI)^2]I_{(p,p)}$$
$$-I[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n](BI_{(q,p)}] = CI_{(m,p)}$$

The next step is to multiply from the right by the identity matrix $I_{(p,q)}$ to get

$$\begin{aligned}
&A[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n]I_{(p,p)}I_{(p,q)}\\
&-I[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n](BI_{(q,p)})I_{(p,q)}\\
&\qquad = CI_{(m,p)}I_{(p,q)}
\end{aligned} \qquad (2.139)$$

Simplifying Eq (2.134) the result obtained is

$$A[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n]I_1$$

$$-I_2[X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n]B = C$$

This implies that

$$X_0+X_1(BI)+X_2(BI)^2+\ldots+X_n(BI)^n$$

is a solution to Eq (2.132), which completes the proof.

Similar results could be achieved if instead of eliminating the Y terms from Eq (2.138), the X terms were eliminated. This is accomplished by multiplying by I, IAI, $I(AI)^2$, ..., $I(AI)^{n+1}$ respectively. The solution generated for Eq (2.132) is

$$Y_0+(IA)Y_1+(IA)^2Y_2+\ldots+(IA)^nY_n$$

Example 2.8: A solution is desired to

$$AXI_1 + I_2XD = E \qquad (2.140)$$

where

$$A = \begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \quad I_2 = \begin{bmatrix}1 & 0\end{bmatrix}, \quad I_1 = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\end{bmatrix},$$

$$D = \begin{bmatrix}1 & 1\end{bmatrix}, \text{ and } E = \begin{bmatrix}0 & -1\\ -1 & -1\end{bmatrix}$$

Thus an equation of the type in Eq (2.132) needs to be solved, and this can be written as

$$\begin{bmatrix}1-\lambda & 0 & 0\\ 0 & -\lambda & 0\end{bmatrix}\left[\begin{bmatrix}x_{11}\\ x_{21}\\ x_{31}\end{bmatrix} + \lambda\begin{bmatrix}x_{12}\\ x_{22}\\ x_{32}\end{bmatrix} + \lambda^2\begin{bmatrix}x_{13}\\ x_{23}\\ x_{33}\end{bmatrix}\right]\begin{bmatrix}1 & 0\end{bmatrix}$$

$$-\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}y_{11} & y_{12} & y_{13}\\ y_{21} & y_{22} & y_{23}\\ y_{31} & y_{32} & y_{33}\end{bmatrix}\begin{bmatrix}1-\lambda & 1\end{bmatrix} = \begin{bmatrix}0 & -1\\ -1 & -1\end{bmatrix} \qquad (2.141)$$

Simplifying this becomes

$$\begin{bmatrix}(1-\lambda)(x_{11}+\lambda x_{12}+\lambda^2 x_{13}) & 0\\ -\lambda(x_{21}+\lambda x_{22}+\lambda^2 x_{23}) & 0\end{bmatrix}$$

$$-\begin{bmatrix}(1-\lambda)(y_{11}+\lambda y_{12}+\lambda^2 y_{13}) & y_{11}+\lambda y_{12}+\lambda^2 y_{13}\\ (1-\lambda)(y_{21}+\lambda y_{22}+\lambda^2 y_{23}) & y_{21}+\lambda y_{22}+\lambda^2 y_{23}\end{bmatrix} = \begin{bmatrix}0 & -1\\ -1 & -1\end{bmatrix} \qquad (2.142)$$

Eq (2.142) can then be written as four equations in six unknowns:

$$\begin{aligned}(1-\lambda)(x_{11}+\lambda x_{12}+\lambda^2 x_{13}) - (1-\lambda)(y_{11}+\lambda y_{12}+\lambda^2 y_{13}) &= 0\\ -\lambda(x_{21}+\lambda x_{22}+\lambda^2 x_{23}) - (1-\lambda)(y_{21}+\lambda y_{22}+\lambda^2 y_{23}) &= -1\\ -\quad (y_{11}+\lambda y_{12}+\lambda^2 y_{13}) &= -1\\ -\quad (y_{21}+\lambda y_{22}+\lambda^2 y_{23}) &= -1\end{aligned} \qquad (2.143)$$

The system of equations in Eq (2.143) can then be quickly solved to get

$$\begin{aligned}y_{11}+\lambda y_{12}+\lambda^2 y_{13} &= 1 \qquad x_{11}+\lambda x_{12}+\lambda^2 x_{13} = 1\\ y_{21}+\lambda y_{22}+\lambda^2 y_{23} &= 1 \qquad x_{21}+\lambda x_{22}+\lambda^2 x_{23} = 1\end{aligned} \qquad (2.144)$$

This implies that the solution vectors are $X = \begin{bmatrix}1\\ 1\\ x\end{bmatrix}$ and $Y = \begin{bmatrix}1\\ 1\\ y\end{bmatrix}$ where x and y are arbitrarily chosen.

Solutions Generated With the Use of Tensor Analysis

This last section attacks the problem of solutions of equations of the type discussed previously in terms of tensors.

In so doing, the work of Lancaster (Ref 21) will be generalized. Before beginning, some basic definitions are needed.

Definition: If $A_{(m,n)}$ and $B_{(p,q)}$ are complex matrices, then the tensor product of A and B written $A \otimes B$, where $A \otimes B$ is a complex matrix, is defined to be the partitioned matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ & \vdots & & \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

The order of $A \otimes B$ is mp x nq.

Definition: Let X be a matrix of order m x n, then X_c is the column vector of order mn x 1, formed by writing all the elements of X in a columnar fashion starting with x_{11} and working across the first row, the second row, and so on until all x_{ij} have been exhausted.

Lancaster has shown (Ref 21:544) that if $A_{(m,m)}$, $B_{(n,n)}$, $I_{1_{(m,m)}}$, and $I_{2_{(n,n)}}$ where $BXI - I_2XI = Q$, then an equivalent expression is $G\vec{x} = \vec{q}$, where $G = (B \otimes I_1^T) - (I_2 \otimes A^T)$. Note that $BXI_1 - I_2XA = Q$ could just as easily have been expressed without the use of the matrices I_1 and I_2.

Example 2.9: Consider the matrix equation

$$AX - XB = Q \tag{2.145}$$

where

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & -2 \end{bmatrix}$$

Eq (2.145) could then be expressed as

$$AXI_1 - I_2XB = Q$$

where

$$I_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then using Lancaster's results a solution to Eq (2.145) by solving

$$[(A \otimes I_1) - (I_2 \otimes B)]\vec{x} = \vec{q} \qquad (2.146)$$

Substituting yields

$$\left[\begin{bmatrix} 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & 1\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ 0\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & 3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{bmatrix} - \begin{bmatrix} 1\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} & 0\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} \\ 0\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} & 1\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} \end{bmatrix}\right]\begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 2 \\ -2 \end{bmatrix} \qquad (2.147)$$

Simplification yields

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 2 \\ -2 \end{bmatrix} \qquad (2.148)$$

Multiplying out Eq (2.148) to form the system of six equations

$$\begin{aligned} x_{11}+x_{21} &= 1 & 2x_{21} &= 4 \\ x_{22} &= 2 & x_{22} &= 2 \\ -x_{12}-x_{13}+x_{23} &= 3 & -x_{22} &= -2 \end{aligned} \qquad (2.149)$$

Solving the system expressed in Eq (2.149), the values are $x_{11} = -1$, $x_{21} = 2$, $x_{22} = 2$, and $-x_{12}-x_{13}+x_{23} = 3$. Thus our choices for x_{12}, x_{13}, and x_{23} can be arbitrary.

From this example another set of solutions could have been generated through the use of generalized inverses.

The last theorem will take the results of Lancaster and extend them to the case where the coefficient matrices are non-square.

Theorem 2.17: Let the matrices A, B, C, D, Q, and X be complex matrices which have the following dimensions: $A_{(m,n)}$, $B_{(p,q)}$, $C_{(m,n)}$, $D_{(p,q)}$, $Q_{(m,p)}$, and $X_{(n,q)}$. If $nq = mp$ and $A^TA \neq 0$, then the matrix equation

$$AXB + CXD = Q \tag{2 150}$$

is equivalent to

$$G\vec{x} = \vec{q} \tag{2.151}$$

where

$$G = A^TA \otimes BB^T + A^TC \otimes BD^T \tag{2.152}$$

Proof: This proof utilizes the results of Lancaster (Ref 21:544) for square matrices. Thus the first operation is to convert Eq (2.150) to one containing square coefficients. Eq (2.150) then becomes

$$\begin{aligned} A^T(AXB)B^T + A^T(CXD)B^T &= A^TQB^T \\ (A^TA)X(BB^T) + (A^TC)X(DB^T) &= A^TQB^T \end{aligned} \tag{2.153}$$

Applying Lancaster, there exists a matrix G defined as

$$G = (A^TA) \otimes (BB^T)^T + (A^TC) \otimes (DB^T)^T$$
$$= A^TA \otimes BB^T + A^TC \otimes BD^T \qquad (2.154)$$

Eq (2.154) is the desired result.

Example 2.10: Consider the equation

$$AXI_1 + I_2XB = Q \qquad (2.155)$$

where

$$A = \begin{bmatrix} 1 & 2 \end{bmatrix}, \; I_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \; I_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}$$
$$B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 \end{bmatrix}, \text{ and } Q = \begin{bmatrix} -4 & 8 & -2 & 4 \end{bmatrix}$$

To find G, the following are needed:

$$A^TA = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}, \; I_1I_1^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; A^TI_2 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix},$$

$$\text{and } I_1B^T = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Thus G is

$$G = A^TA \times I_1I_1^T + A^TI_2 \times I_1B^T \qquad (2.156)$$

Since the equation was changed, the value for Q is also changed, thus

$$A^TQB^T = \begin{bmatrix} -12 & 20 \\ -6 & 10 \end{bmatrix}$$

Solving for the G of Eq (2.156) and thus Eq (2.151):

$$G = \begin{bmatrix} 2\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} & 4\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \\ 1\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} & 2\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \end{bmatrix} + \begin{bmatrix} 2\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix} & 0\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix} \\ 1\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix} & 0\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix} \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix}$$

Eq (2.151) can then be expressed as

$$\begin{bmatrix} 4 & 4 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 \\ 0 & 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} -12 \\ 20 \\ -6 \\ 10 \end{bmatrix} \qquad (2.157)$$

which then implies

$$\begin{aligned} x_{11} + x_{12} + x_{21} &= -3 \\ x_{12} + x_{22} &= 5 \end{aligned} \qquad (2.158)$$

The two equations in Eq (2.158) then imply that any number of arbitrary solutions can be found to satisfy Eq (2.155). Generalized inverses could also have been used to solve this problem.

III. Generalized Inverses and Linear Models

Solutions of the equations being discussed can also be arrived at through the use of generalized inverses. Theorem 2.11 is a case in point. By using this theorem, a whole family of solutions can be generated by use of the general solution

$$X = A^{-}CB^{-} + Z - AA^{-}ZBB^{-} \tag{3.1}$$

where Z is an arbitrary matrix. The first portion of this chapter will deal with more general solutions than the one expressed in Eq (3.1). These results will extend the work of Rao and Mitra (Ref 35).

One advantage that is gained by solving matrix equations through the use of generalized inverses is that a complex system of equations may then be broken down and reduced to a set of more elementary equations. After solving this elementary set of equations, the solution gotten are then joined to find a common solution which is the solution of the original system.

The second half of this chapter deals with applications of the methods developed to the study of linear models. The development begins with a linear model of the form

$$Y = X\beta + U\xi \tag{3.2}$$

where β is an unknown parameter and ξ is a hypothetical random variable with a given dispersion structure but containing unknown variance and covariance components. These two components will then be estimated by use of MINQUE (Minimum Norm Quadratic Unbiased Estimation). After the initial development, more complex models will be studied. This second half extends Rao (Ref 34).

Generalized Inverses

In the evaluation of a matrix equation many solutions are possible. The problem that arises is that some of these solutions may be overlooked. To solve the equation

$$AXB = C \tag{3.3}$$

which is the conclusion of Theorem 2.11, use is made of the solution in Eq (3.1). However, even this general form of solution may omit a vast quantity of solutions. This omission is corrected in the first theorem.

Theorem 3.1: A necessary and sufficient condition for Eq (3.3) to have a solution is that

$$AA^{-}CB^{-}B = C \tag{3.4}$$

in which case the general solution is

$$X = A^{-}CB^{-}+Z-A^{-}AZBB^{-}+(I-A^{-}A)V(I-MM^{-})+(I-N^{-}N)W(I-BB^{-}) \tag{3.5}$$

where M, N, V, W, and Z are arbitrary matrices of appropriate order and I is the identity matrix.

Proof: Let X be an arbitrary matrix such that Eq (3.3) is satisfied. Then if

$$AA^{-}CB^{-}B = AA^{-}CB^{-}B$$

from Eq (3.4)

$$\begin{aligned} AA^{-}(AXB)B^{-}B &= AA^{-}CB^{-}B \\ (AA^{-}A)X(BB^{-}B) &= AA^{-}CB^{-}B \\ AXB &= AA^{-}CB^{-}B \\ C &= AA^{-}CB^{-}B \end{aligned} \tag{3.6}$$

Sufficiency follows by letting X be as expressed in Eq (3.5). Multiplying from the left by A and from the right by B yields

$$\begin{aligned} AXB = {} & AA^{-}CB^{-}B+AZB-AA^{-}AZBB^{-}B+A(I-A^{-}A)V(I-MM^{-})B \\ & + A(I-N^{-}N)W(I-BB^{-})B \end{aligned} \tag{3.7}$$

But

$$AA^{-}AZBB^{-}B = AZB \tag{3.8}$$

and

$$\begin{aligned} A(I-A^{-}A)V(I-MM^{-})B &= (AI-AA^{-}A)V(I-MM^{-})B \\ &= (A-A)V(I-MM^{-})B \\ &= 0 \end{aligned} \tag{3.9}$$

which implies similarly that

$$\begin{aligned} A(I-N^{-}N)W(I-BB^{-})B &= A(I-N^{-}N)W(IB-BB^{-}B) \\ &= A(I-N^{-}N)W(B-B) \\ &= 0 \end{aligned} \tag{3.10}$$

Substituting Eq (3.8), (3.9), and (3.10) into Eq (3.7) yields

$$AXB = AA^{-}CB^{-}B = C \tag{3.11}$$

which implies that Eq (3.5) is a solution of Eq (3.3).

In the sciences it often occurs that a system of matrix equations must be solved in such a way as to yield a common

solution. Thus, if a common solution to the system of equations

$$AX = C$$
$$XB = D \tag{3.12}$$

is needed, one can be found through the applications of the next theorem.

Theorem 3.2 (Rao and Mitra (Ref 35:25)): Let $A_{(m,n)}$, $C_{(m,p)}$, $B_{(p,q)}$, and $D_{(n,q)}$ be given matrices. A necessary and sufficient condition that the consistent system of equations expressed in Eq (3.12) have a common solution is that

$$AD = CB \tag{3.12a}$$

in which case the general expression for a common solution is

$$X = A^{-}C+DB^{-}-A^{-}ADB^{-}+(I-A^{-}A)Z(I-BB^{-}) \tag{3.13}$$

where Z is an arbitrary matrix.

Proof: Let X be a common solution to the system of equations in Eq (3.12) then

$$AX = C \quad \text{and} \quad XB = D$$

which by appropriate multiplications becomes

$$AXB = CB \quad \text{and} \quad AXB = AD \tag{3.14}$$

Thus by setting the two equations of Eq (3.14) equal to each other yields the result wanted in Eq (3.12a).

Sufficiency follows by letting X be expressed as in Eq (3.13). Thus

$$AX = C$$
$$A[A^{-}C+DB^{-}-A^{-}ADB^{-}+(I-A^{-}A)Z(I-BB^{-})] = C \tag{3.15}$$

$$
\begin{aligned}
AA^{-}C+ADB^{-}-AA^{-}ADB^{-}+A(I-A^{-}A)Z(I-BB^{-}) &= C \\
AA^{-}C+ADB^{-}-ADB^{-}+(AI-AA^{-}A)Z(I-BB^{-}) &= C \\
AA^{-}C+(A-A)Z(I-BB^{-}) &= C \\
AA^{-}C &= C
\end{aligned}
\tag{3.15}
$$

The hypothesis of the theorem states that the equations are consistent, thus $AA^{-}C = C$ and X is a solution of the first equation of the system. Similarly,

$$
\begin{aligned}
XB &= D \\
[A^{-}C+DB^{-}-A^{-}ADB^{-}+(I-A^{-}A)Z(I-BB^{-})]B &= D \\
A^{-}CB+DB^{-}B-A^{-}ADB^{-}B+(I-A^{-}A)Z(I-BB^{-})B &= D
\end{aligned}
\tag{3.16}
$$

Making use of Eq (3.12) this last equation can be written as

$$
\begin{aligned}
A^{-}CB+DB^{-}B-A^{-}CBB^{-}B+(I-A^{-}A)Z(IB-BB^{-}B) &= D \\
A^{-}CB+DB^{-}B-A^{-}CB+(I-A^{-}A)Z(B-B) &= D \\
DB^{-}B &= D
\end{aligned}
$$

Again, since the equations are consistent, $DB^{-}B = D$ which implies X is a solution of the second equation of the system. Thus X is the sought-after common solution.

One application of Theorem 3.2 is that it can be used to solve equations of the type

$$AX + XB = E \tag{3.17}$$

To make use of this theorem, the matrix E must first be expressed as a sum of two other matrices. Thus, if

$$E = C + D \tag{3.18}$$

the matrix equation in Eq (3.17) can be written as the system in Eq (3.12) and Theorem 3.2 can be applied. This procedure

can be generalized to solutions of equations of the type

$$AXB + DXE = G \tag{3.19}$$

where $G = C + F$. The method is as in the next theorem.

Theorem 3.3: Let $A_{(p,n)}$, $B_{(n,q)}$, $C_{(p,q)}$, $D_{(s,m)}$, $E_{(n,t)}$, and $F_{(s,t)}$ be given matrices. A necessary and sufficient condition for the system of equations

$$\begin{aligned} AXB &= C \\ DXE &= F \end{aligned} \tag{3.20}$$

to have a common solution is that

$$\begin{aligned} &\text{i)} \quad C = AA^{-}CB^{-}B \\ &\text{ii)} \quad F = DD^{-}FE^{-}E \\ &\text{iii)} \quad A^{-}A(D^{-}FE^{-})BB^{-} = D^{-}D(A^{-}CB^{-})EE^{-} \end{aligned} \tag{3.21}$$

in which case the general expression for a common solution is

$$\begin{aligned} X = {} & A^{-}CB^{-}+D^{-}FE^{-}-A^{-}A(D^{-}FE^{-})BB^{-} \\ & +(I-A^{-}A)V(I-EE^{-})+(I-D^{-}D)W(I-BB^{-}) \end{aligned} \tag{3.22}$$

where V and W are arbitrary matrices.

Proof: From the definition of a generalized inverse it is true that for a matrix H, $HH^{-}H = H$. Thus, if

$$AXB = C$$

it follows that

$$\begin{aligned} (AA^{-}A)X(BB^{-}B) &= C \\ AA^{-}(AXB)B^{-}B &= C \\ AA^{-}CB^{-}B &= C \end{aligned} \tag{3.23}$$

Thus i of Eq (3.21) is shown. A similar argument is used for ii.

$$
\begin{aligned}
DXE &= F \\
(DD^{-}D)X(EE^{-}E) &= F \\
DD^{-}(DXE)E^{-}E &= F \\
DD^{-}FE^{-}E &= F
\end{aligned}
\tag{3.24}
$$

What has been shown thus far is that for i and ii of Eq (3.21) each of the equations of the system in Eq (3.20) has an individual solution X. Condition iii follows from the application of Theorem 3.1. A solution for the first equation of the system in Eq (3.20) is if the arbitrary matrices Z, V and W are the zero matrix,

$$X_1 = A^{-}CB^{-} \tag{3.25}$$

The solution for the second equation of the system is, if V and W are the zero matrix,

$$X_2 = D^{-}FE^{-}+Z-D^{-}DZE^{-}E \tag{3.26}$$

Let X_1 of Eq (3.25) be the Z matrix of Eq (3.26). Thus

$$X_2 = D^{-}FE^{-}+A^{-}CB^{-}-D^{-}DA^{-}CB^{-}E^{-}E \tag{3.27}$$

Now reversing the process, let

$$X_1 = A^{-}CB^{-}+U-A^{-}AUB^{-}B \tag{3.28}$$

and

$$X_2 = D^{-}FE^{-} \tag{3.29}$$

Then let X_2 of Eq (3.29) be the U matrix of Eq (3.28) to yield

$$X_1 = A^{-}CB^{-}+D^{-}FE^{-}-A^{-}AD^{-}FE^{-}B^{-}B \tag{3.30}$$

if the system has a common solution X_1 of Eq (3.30) equals X_2 of Eq (3.27). Hence, $X_1 = X_2$ implies

$$
\begin{aligned}
&= DA^{-}CB^{-}E+DD^{-}FE^{-}E-DD^{-}D(A^{-}CE^{-}EE^{-}E+D(I-A^{-}A)V(I-EE^{-})E \\
&\qquad +D(I-D^{-}D)W(I-BB^{-})E \\
&= DA^{-}CB^{-}E+DD^{-}FE^{-}E-DA^{-}CE^{-}E+D(I-A^{-}A)V(IE-EE^{-}E) \\
&\qquad +(DI-DD^{-}D)W(I-BB^{-})E \\
&= DD^{-}FE^{-}E+D(I-A^{-}A)V(E-E)+(D-D)W(I-BB^{-})E \\
&= DD^{-}FE^{-}E \\
&= F
\end{aligned}
\tag{3.33}
$$

Thus X is also a solution to the second equation of the system in Eq (3.20). This then implies that X is a common solution.

Other conditions can be developed for the system of equations in Eq (3.20) to hold simultaneously. One of these is formalized in the following corollary.

Corollary 3.3.1: A necessary condition for the system of equations given in Eq (3.20) to have a common solution is that the matrices A, B, C, D, E, and F be as defined in Theorem 3.3, i and ii of Eq (3.21) hold and that the matrices C and F may be also defined as follows:

$$
\begin{aligned}
C &= AD^{-}FE^{-}B \\
F &= DA^{-}CB^{-}E
\end{aligned}
\tag{3.34}
$$

Proof: From Theorem 2.11 a solution for AXB = C exists and is

$$X_1 = A^{-}CB^{-}+Z-A^{-}AZBB^{-} \tag{3.35}$$

Also, a solution for DXE = F exists and is

$$X_2 = D^{-}FE^{-}+Y-D^{-}DYEE^{-} \tag{3.36}$$

In Eqs (3.35) and (3.36) the matrices Z and Y are arbitrary.

$$A^{-}CB^{-}+D^{-}FE^{-}-A^{-}AD^{-}FE^{-}B^{-}B = D^{-}FE^{-}+A^{-}CB^{-}-D^{-}DA^{-}CB^{-}E^{-}E \quad (3.31)$$

This can be simplified to

$$A^{-}AD^{-}FE^{-}B^{-}B = D^{-}DA^{-}CB^{-}E^{-}E$$

which is iii of Eq (3.21).

To show that a common solution for X exists, let X be as defined in Eq (3.22) and X will be a common solution if it satisfies both equation of the system in Eq (3.20).

$$\begin{aligned}
AXB &= A[A^{-}CB^{-}+D^{-}FE^{-}-A^{-}A(D^{-}FE^{-})BB^{-}+(I-A^{-}A)V(I-EE^{-}) \\
&\qquad +(I-D^{-}D)W(I-BB^{-})]B \\
&= AA^{-}CB^{-}B+AD^{-}FE^{-}B-AA^{-}A(D^{-}FE^{-})BB^{-}B+A(I-A^{-}A)V(I-EE^{-})B \\
&\qquad +A(I-D^{-}D)W(I-BB^{-})B \\
&= AA^{-}CB^{-}B+AD^{-}FE^{-}B-A(D^{-}FE^{-})B+(AI-AA^{-}A)V(I-EE^{-})B \\
&\qquad +A(I-D^{-}D)W(IB-BB^{-}B) \\
&= AA^{-}CB^{-}B+(A-A)V(I-EE^{-})B+A(I-D^{-}D)W(B-B) \\
&= AA^{-}CB^{-}B \\
&= C
\end{aligned} \quad (3.32)$$

Thus X is a solution for the first equation of the system. Similarly,

$$\begin{aligned}
DXE &= D[A^{-}CB^{-}+D^{-}FE^{-}-A^{-}A(D^{-}FE^{-})BB^{-}+(I-A^{-}A)V(I-EE^{-}) \\
&\qquad +(I-D^{-}D)W(I-BB^{-})]E \\
&= DA^{-}CB^{-}E+DD^{-}FE^{-}E-DA^{-}A(D^{-}FE^{-})BB^{-}E+D(I-A^{-}A)V(I-EE^{-})E \\
&\qquad +D(I-D^{-}D)W(I-BB^{-})E
\end{aligned}$$

Using iii of Eq (3.21), the last equation can be written as

Choose $Z = D\bar{F}E\bar{}$ and $Y = A\bar{C}B\bar{}$, thus the equations become

$$X_1 = A\bar{C}B\bar{}+D\bar{F}E\bar{}-A\bar{A}D\bar{F}E\bar{B}B\bar{} \tag{3.37}$$

and

$$X_2 = D\bar{F}E\bar{}+A\bar{C}B\bar{}-D\bar{D}A\bar{C}B\bar{E}E\bar{}$$

Since a common solution exists from Theorem 3.3

$$X_1 = X_2 \tag{3.38}$$

It then follows that

$$X_1 = D\bar{F}E\bar{} \quad \text{and that} \quad X_2 = A\bar{C}B\bar{} \tag{3.39}$$

Let X_1 be as defined in Eqs (3.37) and (3.39), then

$$\begin{aligned} D\bar{F}E\bar{} &= A\bar{C}B\bar{}+D\bar{F}E\bar{}-A\bar{A}D\bar{F}E\bar{B}B\bar{} \\ 0 &= A\bar{C}B\bar{}-A\bar{A}D\bar{F}E\bar{B}B\bar{} \\ A\bar{C}B\bar{} &= A\bar{A}D\bar{F}E\bar{B}B\bar{} \\ AA\bar{C}B\bar{B} &= AA\bar{A}D\bar{F}E\bar{B}B\bar{B} \\ C &= AD\bar{F}E\bar{B} \end{aligned} \tag{3.40}$$

In going from the fourth to the fifth line of Eq (3.40), use was made of i in Eq (3.21) of Theorem 3.3. To show the second half of Eq (3.34), let X_2 be as defined in Eqs (3.37) and (3.39). Thus

$$\begin{aligned} A\bar{C}B\bar{} &= D\bar{F}E\bar{}+A\bar{C}B\bar{}-D\bar{D}A\bar{C}B\bar{E}E\bar{} \\ 0 &= D\bar{F}E\bar{}-D\bar{D}A\bar{C}B\bar{E}E\bar{} \\ D\bar{F}E\bar{} &= D\bar{D}A\bar{C}B\bar{E}E\bar{} \\ DD\bar{F}E\bar{E} &= DD\bar{D}A\bar{C}B\bar{E}E\bar{E} \\ F &= DA\bar{C}B\bar{E} \end{aligned} \tag{3.41}$$

Similarly use was made of ii in Eq (3.21) of Theorem 3.3, and

the corollary is proved.

Example 3.1: Consider a system of equations as defined in Eq (3.20) where

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 8 & 8 \\ 8 & 8 \\ 8 & 8 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & & & \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ and } F = \begin{bmatrix} 8 \\ 8 \end{bmatrix} \tag{3.42}$$

To see if a common X is possible, the conditions of Corollary 3.3.1 will be checked. Thus,

$$F = DA^{-}CB^{-}E$$

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & & & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-} \begin{bmatrix} 8 & 8 \\ 8 & 8 \\ 8 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ & \\ 1 & 1 \end{bmatrix}^{-} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{3.43}$$

Noble (Ref 30:342) states that if all the elements of an m x n matrix Q are unity, then $A^{-} = (1/mn)A^{T}$. Therefore

$$A^{-} = (1/12) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } B^{-} = (1/4) \begin{bmatrix} 1 & 1 \\ & \\ 1 & 1 \end{bmatrix} \tag{3.44}$$

Thus Eq (3.43) becomes

$$F = \left(\frac{1}{12}\right)\left(\frac{1}{4}\right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ & & & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 8 & 8 \\ 8 & 8 \\ 8 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ & & \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \end{bmatrix} \tag{3.45}$$

Checking the other condition

$$C = AD^{-}FE^{-}B$$

where

$$D^- = (1/8)\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \text{ and } E^- = (1/2)\begin{bmatrix}1 & 1\end{bmatrix}$$

yields

$$C = (\tfrac{1}{8})(\tfrac{1}{2})\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}8\\ 8\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}8 & 8\\ 8 & 8\\ 8 & 8\end{bmatrix}$$

From Theorem 3.3 the value of the common solution X is

$$X = A^-CB^- + D^-FE^- - A^-AD^-FE^-BB^- \tag{3.46}$$

where the arbitrary matrices are zero.

$$A^-CB^- = (\tfrac{1}{12})(\tfrac{1}{4})\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}8 & 8\\ 8 & 8\\ 8 & 8\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \tag{3.47}$$

$$D^-FE^- = (\tfrac{1}{8})(\tfrac{1}{2})\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}8\\ 8\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \tag{3.48}$$

$$A^-AD^-FE^-BB^- = (\tfrac{1}{12})(\tfrac{1}{4})\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

$$= \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \tag{3.49}$$

Making the appropriate substitutions from Eqs (3.47), (3.48), and (3.49) in Eq (3.46) the common solution is

$$X = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} + \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} - \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \tag{3.50}$$

Some applications of Theorem 3.3 and its corollary are to the areas of model reduction and filtering theory. The next example illustrates model reduction.

Example 3.2: Let AXB = C be a model of some phenomenon in which the matrices A, B, and C are defined as

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \text{ and } C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{3.51}$$

Then a possible choice of solution for AXB = C is ${}_aX_b$ where

$${}_aX_b = A^{-}CB^{-} \tag{3.52}$$

Solving for the generalized inverses of A and B yields

$$A^{-} = \left(\frac{2}{73}\right) \begin{bmatrix} 126 & -180 \\ -99 & 204 \\ -120 & 234 \end{bmatrix}$$

$$B^{-} = \left(\frac{1}{10369}\right) \begin{bmatrix} 32976 & -18072 & -23568 & -22860 \\ -45720 & 38640 & 46260 & 43920 \end{bmatrix} \tag{3.53}$$

Thus ${}_aX_b$ can be written as

$${}_aX_b = \left(\frac{2}{73}\right)\left(\frac{1}{10369}\right) \begin{bmatrix} 126 & -180 \\ -99 & 204 \\ -120 & 234 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 32976 & -18072 & -23568 & -22860 \\ -45720 & 38640 & 46260 & 43920 \end{bmatrix}$$

$$= \left(\frac{2}{756937}\right) \begin{bmatrix} 688176 & -1110672 & -1225368 & -1137240 \\ -1338120 & 2159640 & 2382660 & 2211300 \\ -1452816 & 2344752 & 2586888 & 2400840 \end{bmatrix} \tag{3.54}$$

A reduction of this model could be described by the equation DXE = F where the matrices D and E are defined as

$$D = [1 \quad 1/2 \quad 1/3] \quad , \; E = \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \end{bmatrix} \tag{3.55}$$

The matrix F is determined by the equation

$$F = DA^{-}CB^{-}E \tag{3.56}$$

Thus, making the appropriate substitutions

$$F = \left|\begin{matrix} 1 \\ \frac{1}{2} \\ \frac{1}{3} \end{matrix}\right|^{T} \left(\frac{2}{73}\right) \begin{bmatrix} 126 & -180 \\ -90 & 204 \\ -120 & 234 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \left(\frac{1}{10369}\right) \begin{bmatrix} 32976 & -45720 \\ -18072 & 38640 \\ -23568 & 46260 \\ -22860 & 43920 \end{bmatrix}^{T} \begin{bmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{4} \end{bmatrix}$$

$$= \left(\frac{2}{756937}\right)[378468.5]$$

$$= [1] \tag{3.57}$$

The reduced equation DXE = F can then be written as

$$[1 \quad 1/2 \quad 1/3] \; X \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \end{bmatrix} = [1] \tag{3.57}$$

To solve Eq (3.57) the generalized inverses of D and E are needed.

$$D^{-} = \left(\frac{1}{49}\right)\begin{bmatrix} 36 \\ 18 \\ 12 \end{bmatrix} \text{ and } E^{-} = \left(\frac{1}{205}\right)[144 \quad 72 \quad 48 \quad 36] \tag{3.58}$$

Solving Eq (3.57) for the matrix X yields

$$d^{X}e = \left(\frac{1}{49}\right)\begin{bmatrix} 36 \\ 18 \\ 12 \end{bmatrix}[1] \left(\frac{1}{205}\right)[144 \quad 72 \quad 48 \quad 36] \tag{3.59}$$

$$= \left(\frac{1}{10045}\right) \begin{bmatrix} 5148 & 2592 & 1728 & 1296 \\ 2592 & 1296 & 864 & 648 \\ 1728 & 864 & 576 & 432 \end{bmatrix}$$

Eq (3.59) is a solution for the reduced equation. If there existed interest in the common solution, it could be found by application of Eq (3.22). For purposes of this example, let the arbitrary matrices V and W be the zero matrix. Thus, the common solution is

$$X = A^{-}CB^{-}+D^{-}FE^{-}-A^{-}A(D^{-}FE^{-})BB^{-} \tag{3.60}$$

This could also be expressed as

$$X = {}_{a}X_{b} + {}_{d}X_{e} - A^{-}A_{d}X_{e}BB^{-} \tag{3.61}$$

The values of ${}_{a}X_{b}$ and ${}_{d}X_{e}$ are known for Eqs (3.54) and (3.59). What needs to be found is the value of

$$A^{-}A_{d}X_{e}BB^{-} \tag{3.62}$$

$$A^{-}A = \left(\frac{1}{73}\right)\begin{bmatrix} 72 & 6 & -6 \\ 6 & 37 & 36 \\ -6 & 36 & 37 \end{bmatrix}$$

and

$$BB^{-} = \left(\frac{1}{10369}\right)\begin{bmatrix} 10116 & 1248 & -438 & -900 \\ 1248 & 3844 & 3636 & 3210 \\ -438 & 3636 & 3709 & 3360 \\ -900 & 3210 & 3360 & 6429 \end{bmatrix} \tag{3.63}$$

Thus the value of Eq (3.62) is the product of the matrices in Eq (3.63) and Eq (3.59). The result then is

$$A^{-}A_{d}X_{e}BB^{-} = \left(\frac{1}{10045}\right)\begin{bmatrix} 2592 & 1296 & 864 & 648 \\ 1296 & 643 & 432 & 324 \\ 864 & 432 & 288 & 216 \end{bmatrix} = {}_{d}X_{e} \tag{3.64}$$

This then implies that the common solution expressed in Eq (3.61) is

$$X = {}_aX_b + {}_dX_e - {}_dX_e$$

$$X = {}_aX_b \tag{3.65}$$

and the value for ${}_aX_b$ is given in Eq (3.54).

Solutions to system with more equations can also be solved by applications of the above methods. These solutions are based upon appropriate choices of the arbitrary matrices.

Theorem 3.4: Let the matrices A, B, C, D, E, F, G, H, and K be of appropriate dimension. A necessary and sufficient condition that the system of equations

$$\begin{aligned} AXB &= C \\ DXE &= F \\ GXH &= K \end{aligned} \tag{3.66}$$

to have a common solution is that

$$\begin{aligned} &\text{i)} && C = AA^{-}CB^{-}B \\ &\text{ii)} && F = DD^{-}FE^{-}E \\ &\text{iii)} && K = GG^{-}KH^{-}H \\ &\text{iv)} && C = AD^{-}FE^{-}B \\ &\text{v)} && G^{-}KH^{-} = A^{-}AG^{-}KH^{-}BB^{-} = D^{-}FE^{-} \end{aligned} \tag{3.67}$$

in which case the general expression for a common solution is

$$X = A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-} \tag{3.68}$$

Proof: The proof of i and ii follow from Eqs (3.23) (3.24). To show iii, let GXH = K. Then from the definition of generalized inverse it follows that

$$(GG^{-}G)X(HH^{-}H) = K$$
$$GG^{-}(GXH)H^{-}H = K \qquad (3.69)$$
$$GG^{-}XH^{-}H = K$$

This then proves iii. Section iv of Eq (3.67) follows as in Eq (3.40) of the Corollary to Theorem 3.3. Part v of Eq (3.67) follows from an application of Theorem 3.3. A solution for DXE = F and GXH = K exists and is

$$X_1 = G^{-}KH^{-}+D^{-}FE^{-} \qquad (3.70)$$

where the arbitrary matrices are taken as the zero matrix.

From Theorem 2.11 a solution for AXB = C exists and is

$$X_2 = A^{-}CB^{-}+Q-A^{-}AQBB^{-} \qquad (3.71)$$

Let the Q of Eq (3.71) be the value of X_1 from Eq (3.70). Then

$$X_2 = A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-} \qquad (3.72)$$

Eq (3.72) is the common solution of the system in Eq (3.66) if it exists, and $X_1 = X_2$ which implies that X_2 is a solution of $DX_2E = F$.

$$\begin{aligned} DX_2E &= DA^{-}CB^{-}E+DD^{-}FE^{-}E+DG^{-}KH^{-}E-DA^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-}E \\ &= DA^{-}CB^{-}E+F+DG^{-}KH^{-}E-DA^{-}AD^{-}FE^{-}BB^{-}E-DA^{-}AG^{-}KH^{-}BB^{-}E \\ &= DA^{-}CB^{-}E+F+DG^{-}KH^{-}E-DA^{-}CB^{-}E-DA^{-}AG^{-}KH^{-}BB^{-}E \\ &= F+DG^{-}KH^{-}E-DA^{-}AG^{-}KH^{-}BB^{-}E \end{aligned} \qquad (3.73)$$

But $DX_2E = F$, thus Eq (3.73) becomes

$$F = F+DG^{-}KH^{-}E-DA^{-}AG^{-}KH^{-}BB^{-}E$$

which implies

$$DG^{-}KH^{-}E = DA^{-}AG^{-}KH^{-}BB^{-}E$$
$$G^{-}KH^{-} = A^{-}AG^{-}KH^{-}BB^{-} \tag{3.74}$$

Also, since Eq (3.72) is a common solution $GX_2H = K$. Thus

$$\begin{aligned} K &= G[A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-}]H \\ &= GA^{-}CB^{-}H+GD^{-}FE^{-}H+GG^{-}KH^{-}H-GA^{-}AD^{-}FE^{-}BB^{-}H-GA^{-}AG^{-}KH^{-}BB^{-}H \\ &= GA^{-}CB^{-}H+GD^{-}FE^{-}H+K-GA^{-}CB^{-}H-GA^{-}AG^{-}KH^{-}BB^{-}H \\ 0 &= GD^{-}FE^{-}H-GA^{-}AG^{-}KH^{-}BB^{-}H \end{aligned} \tag{3.75}$$

$$GD^{-}FE^{-}H = GA^{-}AG^{-}KH^{-}BB^{-}H$$
$$D^{-}FE = A^{-}AG^{-}KH^{-}BB^{-}$$

The conclusions from Eqs (3.74) and (3.75) then imply v of Eq (3.67). The proof of sufficiency follows by letting X as defined in Eq (3.68) be a common solution for the system of equations in Eq (3.66). Thus X must satisfy each of these equations.

$$\begin{aligned} AXB &= A[A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-}]B \\ &= AA^{-}CB^{-}B+AD^{-}FE^{-}B+AG^{-}KH^{-}B-AA^{-}AD^{-}FE^{-}BB^{-}B-AA^{-}AG^{-}KH^{-}BB^{-}B \\ &= AA^{-}CB^{-}B+AD^{-}FE^{-}B+AG^{-}KH^{-}B-AD^{-}FE^{-}B-AG^{-}KH^{-}B \\ &= AA^{-}CB^{-}B \\ &= C \end{aligned} \tag{3.76}$$

This last statement is from hypothesis i, Eq (3.67). X is, therefore, a solution to the first equation of the system. Solving the second equation of the system is next.

$$\begin{aligned} DXE &= D^{-}[A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-}]E \\ &= DA^{-}CB^{-}E+DD^{-}FE^{-}E+DG^{-}KH^{-}E-DA^{-}AD^{-}FE^{-}BB^{-}E-DA^{-}AG^{-}KH^{-}BB^{-}E \end{aligned}$$

Making use of hypothesis ii, iv, and v of Eq (3.67) yields

$$DXE = DA^{-}CB^{-}E+F+DG^{-}KH^{-}E-DA^{-}CB^{-}E-DG^{-}KH^{-}E = F$$

Following a similar line of reasoning, it can easily be shown that

$$GXH = G[A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-A^{-}A(D^{-}FE^{-}+G^{-}KH^{-})BB^{-}]H$$
$$= GA^{-}CB^{-}H+GD^{-}FE^{-}H+GG^{-}KH^{-}H-GA^{-}AD^{-}FE^{-}BB^{-}H-GA^{-}AG^{-}KH^{-}BB^{-}H$$

Making use of hypothesis iii, iv and v of Eq (3.67) yields

$$GXH = GA^{-}CB^{-}H+GD^{-}FE^{-}H+K-GA^{-}CB^{-}H-GD^{-}FE^{-}H$$
$$= K$$

Thus X is a common solution and the theorem is proved.

Corollary 3.4.1: Let the matrices A,B,C,D,E,F,G,H, and K be of appropriate dimension. A necessary and sufficient condition that the system of equations given in Eq (3.66) have a common solution is that items i, ii, iii of Eq (3.67) hold and that

$$\begin{aligned} &\text{vi)} \quad F = DG^{-}KH^{-}E \\ &\text{vii)} \quad G^{-}KH^{-} = D^{-}DA^{-}CB^{-}EE^{-} = A^{-}CB^{-} \end{aligned} \tag{3.77}$$

are also true. A common solution then exists and is

$$X = A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-D^{-}D(A^{-}CB^{-}+G^{-}KH^{-})EE^{-} \tag{3.78}$$

Proof: The proof is identical to that of the theorem with only variable name changes.

Corollary 3.4.2: Let the matrices A,B,C,D,E,F,G,H, and K be of appropriate dimension. A necessary and sufficient condition that the system of equations given in Eq (3.66) have a common solution is that item i, ii, iii of Eq (3.67) hold and that

$$\text{viii)}\quad K = GA^{-}CB^{-}H$$
$$\text{ix)}\quad D^{-}FE^{-} = G^{-}GD^{-}FE^{-}HH^{-} = A^{-}CB^{-} \tag{3.79}$$

are also true. A common solution then exists and is

$$X = A^{-}CB^{-}+D^{-}FE^{-}+G^{-}KH^{-}-G^{-}G(A^{-}CB^{-}+D^{-}FE^{-})HH^{-} \tag{3.80}$$

Proof: A proof is identical to that of the theorem with only variable name changes.

Example 3.3: Solve for the common solution of the system of equations as given in Eq (3.66) where A, B, C, D, E, and F are as given in Eq (3.42). Let G, H, and K be defined as follows

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad K = \begin{bmatrix} 8 & 8 & 8 \\ 8 & 8 & 8 \\ 8 & 8 & 8 \\ 8 & 8 & 8 \end{bmatrix} \tag{3.82}$$

To see if a common solution is possible, the conditions of the theorem must be checked. First find the necessary generalized inverses. A^{-} and B^{-} are as given in Eq (3.45) and the other inverses will be found by using the method of Noble (Ref 30:342), therefore

$$D^{-} = \left(\frac{1}{8}\right)\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad E^{-} = \left(\frac{1}{2}\right)\begin{bmatrix} 1 & 1 \end{bmatrix}, \quad G^{-} = \left(\frac{1}{16}\right)\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\text{and } H^{-} = \left(\frac{1}{6}\right)\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{3.83}$$

Condition i.

$$C = AA^{-}CB^{-}B$$

$$= \left(\frac{1}{12}\right)\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}8 & 8\\ 8 & 8\\ 8 & 8\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\\ & & & \end{bmatrix}\left(\frac{1}{4}\right)$$

$$= \begin{bmatrix}8 & 8\\ 8 & 8\\ 8 & 8\end{bmatrix}$$

Condition ii.

$F = DD^{-}FE^{-}E$

$$= \left(\frac{1}{8}\right)\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}8\\ 8\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}\left(\frac{1}{2}\right) = \begin{bmatrix}8\\ 8\end{bmatrix}$$

Condition iii.

$K = GG^{-}KH^{-}H$

$$= \left(\frac{1}{16}\right)\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\ & & \end{bmatrix}\left(\frac{1}{6}\right)$$

$$= \begin{bmatrix}8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\end{bmatrix}$$

Condition iv.

$C = AD^{-}FE^{-}B$

$$= \left(\frac{1}{8}\right)\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}8\\ 8\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\left(\frac{1}{2}\right)$$

$$= \begin{bmatrix}8 & 8\\ 8 & 8\\ 8 & 8\end{bmatrix}$$

Condition v.

$G^{-}KH^{-} = D^{-}FE^{-}$

$D^{-}FE^{-}$ was evaluated in Eq (3.48) and is equal to $\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}$.

$$G^{-}KH^{-} = \left(\frac{1}{16}\right)\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\\ 8 & 8 & 8\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\left(\frac{1}{6}\right)$$

$$= \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \qquad (3.84)$$

Thus all the conditions of the theorem are met and the common solution is given by Eq (3.68). Making substitutions from Eqs (3.47), (3.48), and (3.84) yields

$$X = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} + \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} + \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} -$$

$$\left(\frac{1}{12}\right)\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\left(\begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} + \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix}\right)\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\left(\frac{1}{4}\right)$$

$$= \begin{bmatrix}3 & 3\\ 3 & 3\\ 3 & 3\\ 3 & 3\end{bmatrix} - \left(\frac{1}{48}\right)\begin{bmatrix}3 & 3 & 3 & 3\\ 3 & 3 & 3 & 3\\ 3 & 3 & 3 & 3\\ 3 & 3 & 3 & 3\end{bmatrix}\begin{bmatrix}2 & 2\\ 2 & 2\\ 2 & 2\\ 2 & 2\end{bmatrix}\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$$

$$= \begin{bmatrix}3 & 3\\ 3 & 3\\ 3 & 3\\ 3 & 3\end{bmatrix} - \begin{bmatrix}2 & 2\\ 2 & 2\\ 2 & 2\\ 2 & 2\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \qquad (3.85)$$

Thus a common solution exists and is equal to the last matrix of Eq (3.85).

Linear Modeling

The theory developed in the section on generalized inverses will be used to help in the estimation of some of the

components of linear models. Variance and estimation error have been studied with regard to Reduced Order Filters by Asher (Ref 1) and the estimation of both variance and covariance have been considered by Rao (Ref 34).

A linear model is one that can be expressed as

$$Y = X\beta + U_1\xi_1 + \ldots + U_k\xi_k \tag{3.86}$$

where Y is an n-vector of random variables, X is a given n x m matrix, β is an m-vector of unknown parameters, the U_i's are given n x c_i matrices, and the ξ_i is a c_1-vector of uncorrelated random variables with a zero mean value and a dispersion matrix $\sigma_i^2 I_{c_i}$, i = 1,...k where the variances are unknown.

An alternate method of expressing Eq (3.86) is as

$$Y = X\beta + U\xi \tag{3.87}$$

where $U = (U_1; U_2; \ldots; U_k)$ and $\xi^T = (\xi_1^T; \xi_2^T; \ldots; \xi_k^T)$.

To estimate the variance components of the linear function

$$p_1\sigma_1^2 + \ldots + p_k\sigma_k^2 \tag{3.88}$$

the quadratic function Y^TAY of the random variable Y in Eq (3.86) or (3.87) will be used. To find this matrix A, some criteria will need to be developed. First the translation of the β parameter should be invariant. Thus consider Eq (3.87) which can be written as

$$Y - X\beta_0 = X\beta - X\beta_0 + U\xi \tag{3.89}$$

Let $X\beta - X\beta_0 = X\gamma$, then the estimator of Eq (3.88) becomes

$$(Y - X\beta_0)^T A(Y - X\beta_0) \tag{3.90}$$

Since the variance should be as close to zero as possible, Eq (3.90) should be set equal to zero. Expanding yields

$$Y^T AY - (X\beta_0)^T AY - Y^T AX_0 - (X\beta_0)^T AX_0 = 0 \tag{3.91}$$

If β is to remain invariant under translation, then from Eq (3.91) $AX = 0$.

A second criteria is that the estimate be unbiased. Using the restriction that $AX = 0$, the estimate can be expressed as

$$Y^T AY = \xi^T U^T AU\xi \tag{3.92}$$

which is an expression in terms of the hypothetical vector variable ξ. If Eq (3.92) is unbiased for Eq (3.88), for all σ_i^2, then

$$E(\xi^T U^T AU\xi) = \sum_{i=1}^{k} E(\xi_i^T U_i^T AU_i \xi_i) = \sum_{i=1}^{k} \sigma_i^2 \text{ tr } U_i^T AU_i \tag{3.93}$$

However Eq (3.93) is another expression for Eq (3.88). Thus this implies that

$$\text{tr } U_i^T AU_i = p_i \tag{3.94}$$

for all $i = 1, \ldots, k$.

The third criteria is that of minimum norm. This says that if the hypothetical variable ξ were known, then a natural estimator of Eq (3.88) is

$$(p_1/c_1)\xi_1^T\xi_1 + \ldots + (p_k/c_k)\xi_k^T\xi_k = \xi^T D \qquad (3.95)$$

where D is an arbitrary diagonal matrix. Thus there presently exists two estimators, the estimator of Eq (3.92) and the estimator of Eq (3.95). Taking the difference yields

$$\xi^T U^T A U \xi - \xi^T D \xi = \xi^T (U^T A U - D) \xi \qquad (3.96)$$

This difference can then be made small, in some sense, by taking the norm of the matrix

$$\| U^T A U - D \| \qquad (3.97)$$

The norm of Eq (3.97) can be any acceptable norm that satisfies the properties of a norm (Ref 14:198). One choice of norm is the Euclidean norm defined as

$$\| U^T A U - D \|^2 = \text{tr } (U^T A U - D)(U^T U A - D) \qquad (3.98)$$

Thus the problem of finding the Minimum Norm Quadratic Unbiased Estimator (MINQUE) of Eq (3.95) is one of finding a matrix A such that Eq (3.98) is a minimum subject to the conditions

$$\begin{aligned} AX &= 0 \\ \text{tr } AV_i &= p_i \ , \ i = 1, \ldots k \end{aligned} \qquad (3.99)$$

where $V_i = U_i^T U_i$.

With these concepts in mind, consider the model given in Eq (3.86), where X is a given m x n matrix and β is an m-vector of unknown parameter. ξ_i is a q-vector such that

$$E(\xi_i) = 0 \ , \ E(\xi_i \xi_i^T) = S \ , \ \text{Cov}(\xi_i, \xi_j) = 0, \ i \neq j \qquad (3.100)$$

The problem that is now considered is one of estimating the $q(q+1)/2$ components of the symmetric matrix S given in Eq (3.100) or one of estimating the linear function of S when the vector is unknown. The dispersion of Eq (3.86) is given by

$$D(Y) = U_1 S U_1^T + \ldots + U_k S U_k^T \tag{3.101}$$

The problem of estimating the given linear function of the elements in S, which can be written as

$$\text{tr } SQ, \tag{3.102}$$

where Q is an arbitrary symmetric matrix, can be solved by letting $Y^T AY$ be an unbiased quadratic estimate of Eq (3.102) with the restriction that $AX = 0$. Thus

$$E(Y^T AY) = \text{tr } AD(Y) = \text{tr } S(U_1^T A U_1 + \ldots + U_k^T A U_k) \tag{3.103}$$

Comparing Eqs (3.102) and (3.103), it is obvious that

$$Q = U_1^T A U_1 + \ldots + U_k^T A U_k \tag{3.104}$$

If the $\xi_1, \ldots, \xi_k$ are known, then a natural choice for estimator of S is

$$(1/k)(\xi_1^T \xi_1 + \ldots + \xi_k^T \xi_k) \tag{3.105}$$

and a natural choice for the estimator of tr SQ is

$$\text{tr } (1/k)(\xi_1^T \xi_1 + \ldots + \xi_k^T \xi_k)Q = (1/k)(\xi_1^T Q_1 + \ldots + \xi_k^T Q_k) \tag{3.106}$$

Now the estimator that was initially proposed is

$$Y^T AY = (U_1 \xi_1 + \ldots + U_k \xi_k)^T A (U_1 \xi_1 + \ldots + U_k \xi_k) \tag{3.107}$$

Considering the estimators in Eqs (3.106) and (3.107) as quadratic expressions in ξ, then what remains is to minimize the norm of their difference. Finally, the problem is one of minimizing

$$\|U^T A U\| \tag{3.108}$$

subject to the conditions

$$AX = 0$$

$$\sum_{i=1}^{k} U_i^T A U_i = Q \tag{3.109}$$

The preceding development is due to Rao (Ref 34). The methods of the preceding section are utilized in the finding of the matrix A in Eqs (3.108) and (3.109).

Example 3.4: Consider a linear model of the type in Eq (3.87) where Y is an n-vector of random variables, β is an m-vector of unknown parameters, and the matrices X, U, and ξ are defined as follows

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\xi^T = [1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1] \tag{3.110}$$

The matrix U can be expressed as

$$U = [U_1 : U_2 : U_3] = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

and

$$\xi^T = [\xi_1^T : \xi_2^T : \xi_3^T] = [(1\ 0\ 0)^T : (0\ 1\ 0)^T : (0\ 0\ 1)^T]$$

The problem at hand is to solve Eq (3.108) subject to the conditions of Eq (3.109). Thus the equations to be solved are

$$A \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = 0 \tag{3.111}$$

and

$$\sum_{i=1}^{3} U^T A U = Q$$

Since Q is an arbitrarily chosen matrix, let Q be defined as

$$Q = \begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$$

The procedures of the last section will be employed to solve the above system of equations. From Penrose (Ref 31) the equation AX = 0 can be solved by use of the formula

$$A = cX^{-} + W(XX^{-} - I) \tag{3.112}$$

where c is a constant matrix and W is arbitrary.

Again applying Noble (Ref 30:342) to find the inverse of X, the generalized inverse of X is

$$X^{-} = \left(\frac{1}{12}\right) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{3.113}$$

Solving Eq (3.112) yields

$$A_x = 0X^{-} + W\left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \left(\frac{1}{12}\right) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

$$= W\left(\left(\frac{1}{3}\right)\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} - \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\right)$$

$$= W\left(\left(\frac{1}{3}\right)\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}\right) \qquad (3.114)$$

Since W is arbitrary, let W be the identity. Thus A is

$$A_x = \left(\frac{1}{3}\right)\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix} \qquad (3.115)$$

The next step is to solve

$$\sum_{i=i}^{3} U^T A U = Q \qquad (3.116)$$

Since each of the U_i are the same, the equation above could be written as

$$3U^T A U = Q \qquad (3.117)$$

Applying Theorem 2.11 to solve for A yields

$$A = (1/3)U^{T-}QU^T + Z - U^{T-}U^T Z U U^- \qquad (3.118)$$

Notice that $U = U^T$, thus $U^- = U^{T-}$ and this value is

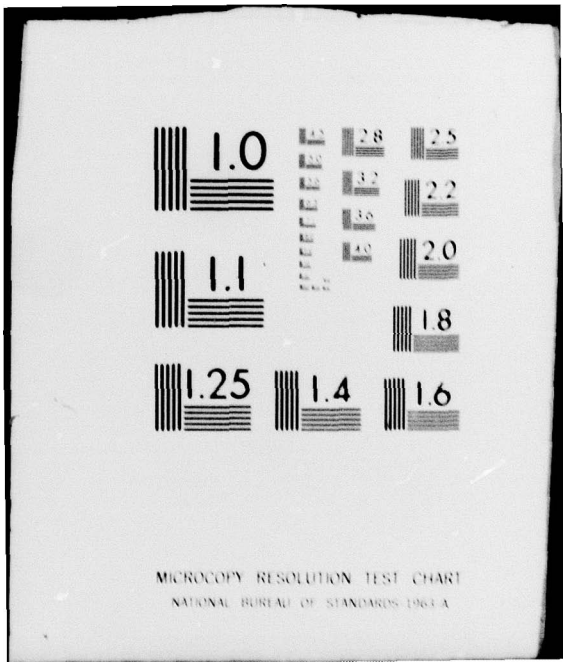
$$U^- = U^{T-} = (1/4)\begin{bmatrix}1 & 0 & 1\\ 0 & 4 & 0\\ 1 & 0 & 1\end{bmatrix} \qquad (3.119)$$

In Eq (3.118), since the end result is a common solution, let the arbitrary matrix Z equal A_x. Solving Eq (3.118) yields

AD-A080 216 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOO--ETC F/G 12/1
SOLUTIONS OF THE MATRIX EQUATION AXB + CXD = E.(U)
DEC 79 T S POMYKALSKI
UNCLASSIFIED AFIT/GOR/MA/79D-6 NL

2 OF 2
AD
A080216

END
DATE
FILMED
3-80
DDC


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

$$A = (\tfrac{1}{3})(\tfrac{1}{4})(\tfrac{1}{4})\begin{bmatrix}1 & 0 & 1\\ 0 & 4 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}-2 & 2. & -2\\ 2 & -2 & 2\\ -2 & 2 & -2\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ 0 & 4 & 0\\ 1 & 0 & 1\end{bmatrix} + (\tfrac{1}{3})\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}$$

$$- (\tfrac{1}{3})(\tfrac{1}{4})(\tfrac{1}{4})\begin{bmatrix}2 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 2\end{bmatrix}\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}2 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 2\end{bmatrix} \qquad (3.120)$$

In Eq (3.120) the quantity

$$(\tfrac{1}{4})\begin{bmatrix}2 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 2\end{bmatrix}$$

is the product of $U^{T-}U^{T} = UU^{-}$. Eq (3.120) simplifies to

$$A = (\tfrac{1}{48})\begin{bmatrix}-8 & 16 & -8\\ 16 & -32 & 16\\ -8 & 16 & -8\end{bmatrix} + (\tfrac{1}{3})\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}$$

$$- (\tfrac{1}{48})\begin{bmatrix}-8 & 16 & -8\\ 16 & -32 & 16\\ -8 & 16 & -8\end{bmatrix} \qquad (3.121)$$

$$A = (\tfrac{1}{3})\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}$$

Thus a common solution has been arrived at. With the value of A now known, it is a simple matter to minimize the quantity $\|U^{T}AU\|$.

IV. Conclusions and Recommendations

The increasing use of matrix equations in the engineering sciences has stimulated a rapidly growing interest in how to best solve these equations. The techniques developed in this thesis can, depending on the equation, be tedious to do by hand, but all can be coded for computer application. Several different methods of solution have also been presented so that if one method fails to yield acceptable results, another way may be implemented that will in turn be satisfactory.

Additional attention can be paid to the area of quadratic matrix equations. Some of the techniques that have been discussed may prove to be of value in the solution of this type equation.

Bibliography

1. Asher, R.B., K.D. Herring and J.C. Rylea. "Bias, Variance, and Estimation Error in Reduced Order Filters," Automatica, 12: 5-9-600 (1976).

2. Bar-Ness, Y., and G. Langholz. "The Solutions of the Matrix Equation XC-BX=D as an Eigen Value Problem," International Journal of Systems Science, 8: 385-396 (1977)

3. Ben-Israel, Adi, and T.N.E. Greville. Generalized Inverses: Theory and Applications. New York: John Wiley & Sons, Inc., 1974.

4. Browne, E.T. Introduction to the Theory of Determinants and Matrices. Chapel Hill, N.C.: University of North Carolina Press, 1958.

5. Cohen, A.M. et al. Numerical Analysis. New York: John Wiley & Sons, Inc., 1973.

6. Dickson, L.E. New First Course in the Theory of Equations. New York: John Wiley & Sons, Inc., 1939.

7. Duff, I.S. "On Permutations to Block Triangular Form," Journal Instruction Mathematics Applications, 19: 339-342 (1977).

8. Friedman, B. Principles and Techniques of Applied Mathematics. New York: John Wiley & Sons, Inc., 1956.

9. Frost, M.G. and C. Story. "Equivalence of a Matrix Over R[s,z] With its Smith Form," International Journal of Control, 28: 665-671 (1978).

10. Gantmacher, F.R. The Theory of Matrices, Vol. II. New York: Chelsea Publishing Co., 1959.

11. Herstein, I.N. Topics in Algebra, 2d ed. Lexington, Mass.: Xerox College Publishing, 1975.

12. Hoskins, W.D. and D.J. Walton. "The Numerical Solution of the Time-Dependent Matrix Equation A(+)V(+)+WA(t)-G(t)," International Journal of Systems Science, 9: 1019-1028 (1978).

13. Jennings, Walter. First Course in Numerical Methods. New York: The MacMillan Company, 1964.

14. Jones, John Jr. "Solutions of Certain Matrix Equations," Proceedings of the American Mathematical Society, 31: 333-339 (1972).

15. Jones, J. and T.S. Pomykalski. "The Matrix Equation AXB+CXD = E," Notices of the American Mathematical Society, 26: A-457 (1979).

16. Jones, J., T.S. Pomykalski and D.W. Repperger. "Reducibility of Algebraic Systems of Equations," Notices of the American Mathematical Society, publication pending.

17. Khatri, C.G. and S.K. Mitra. "Hermitian and Non-Negative Definite Solutions of Linear Matrix Equations," Siam Journal of Applied Mathematics, 31: 579-585 (1976).

18. Kitagawa, G. "An Algorithm for Solving the Matrix Equation $X = FXF^T + S$," International Journal of Control, 25: 745-753 (1977).

19. Komkov, Vadim. "A Generalization of Leighton's Variational Theorem," Applicable Analysis, 1: 377-383 (1972).

20. Kucera, V. "The Matrix Equation AX+XB=C," Siam Journal of Applied Mathematics, 26: 15-25 (1974).

21. Lancaster, Peter. "Explicit Solutions of Linear Matrix Equations," Siam Review, 12: 544-566 (1977)

22. _____. Theory of Matrices. New York: Academic Press, 1969.

23. Leuthauser, J.L. Solutions of Certain Matrix Equations with Applications to Engineering Sciences. Master's Thesis. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, 1976. AD A032667.

24. Lovass-Nagy, V. and D.L. Powers. "A Note on Block-Diagonalization of Some Partitioned Matrices," Linear Algebra and Its Applications, 5: 339-346 (1972).

25. _____. "On Least Squares Solutions of an Inconsistent Singular Equation AX+XB=C," Siam Journal of Applied Mathematics, 31: 84-88 (1976).

26. Ma, Er-Chieh. "A Finite Series Solution of the Matrix Equation AX-XB = C," Siam Journal of Applied Mathematics, 14: 490-495 (1966).

27. Martensson, K. "On the Matrix Riccati Equation," Information Sciences, 3: 17-49 (1971).

28. Mitra, Sujit Kumar. "The Matrix Equation AXB+CXD = E," Siam Journal of Applied Mathematics, 32: 823-825 (1977).

29. Nering, Evan D. Linear Algebra and Matrix Theory. 2d ed. New York: John Wiley & Sons, Inc., 1963.

30. Noble, Ben and James W. Daniel. Applied Linear Algebra. 2d ed. Englewood Cliffs, N.J: Prentice-Hall, Inc., (1977).

31. Penrose, R. "A Generalized Inverse for Matrices," Proceedings Cambridge Philosophical Society, 51: 406-413 (1955).

32. Potter, James E. "Matrix Quadratic Solutions," Siam Journal of Applied Mathematics, 14: 49--501 (1966).

33. Ramachandran, Vijaya. "Exact Reduction of a Polynomial to the Smith Normal Form," IEEE Transactions on Automatic Control, 24: 638-641 (1979).

34. Rao, C. Radhakrishna. "Estimation of Variance and Covariance Components in Linear Models," Journal of the American Statistical Association, 67: 112-115 (1972).

35. _____ and S.K. Mitra. Generalized Inverses of Matrices and Its Applications. New York: John Wiley & Sons, Inc., 1971.

36. Rosenblum, Marvin. "On the Operator Equation BX-XA=Q," Duke Mathematical Journal, 23: 263-270 (1956).

37. Rosenbrock, H.H. and J. Storey. Mathematics of Dynamical Systems. New York: John Wiley & Sons, Inc., 1970.

38. Roth, William E. "The Equations AX-YB=C and AX-XB=C in Matrices," Proceedings of the American Mathematical Society, 3: 392-396 (1952).

39. Stein, Junior. Conjugate Direction Algorithms in Numerical Analysis and Optimization. Toledo: University of Toledo, 1975. NSF GP 40175.

40. Stoyle, P.N.R. and A.I.G. Vardulakis. "The Mechanism of Decoupling," International Journal of Control, 29: 589-605 (1979).

41. Vetter, W.J. "Vector Structures and Solutions of Linear Matrix Equations," Linear Algebra and Its Applications, 10: 181-188 (1975).

42. Wimmer, H.K. "A Jordan Factorization Theorem for Polynomial Matrices," Proceedings of the American Mathematical Society, 75: 201-206 (1979).

43. Wolovich, W.A. "Skew Prime Polynomial Matrices," IEEE Transactions on Automatic Control, 23: 880-887, (1978).

44. Wu, Min-Yen, I.M. Horowitz and J.C. Dennison. "On Solution, Stability, and Transformation of Linear Time-Varying Systems," International Journal of Control, 22: 169-180 (1975).

45. _____. "Transformation of a Linear Time-Varying System Into a Linear Time-Invariant System," International Journal of Control, 27: 589-602 (1978).

46. Ziedan, I.E. "Explicit Solution of the Liapanov Matrix Equation," IEEE Transactions on Automatic Control, 17: 379-381 (1972).

Vita

Captain Thomas S. Pomykalski was born in Chicago, Illinois in 1947. After graduating from St. Rita's High School there, he attended St. Mary's College, Winona, Minnesota from which he received a B.A. degree in mathematics in 1969. After teaching at Aquinas High School, La Crosse, Wisconsin for two years, he joined the Air Force in 1971. Subsequent assignments included Navigator and Electronic Warfare Officer Schools at Mather AFB, California. In 1973 he was assigned to the 16th Special Operations Squadron, SEA, where he served as an Instructor and Evaluator Electronic Warfare Officer in the AC 130 A/E/H. Upon returning to the United States, he was assigned to the 46th Bombardment Squadron, Grand Forks AFB, North Dakota, where he was an Instructor Electronic Warfare Officer until his assignment to AFIT in 1978.

Permanent Address: Box 256
Route 5
Hayward, WI 54843

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/GOR/MA/79D-6	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SOLUTIONS OF THE MATRIX EQUATION AXB + CXD = E		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Thomas S. Pomykalski Capt USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT) Wright-Patterson AFB, OH 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 1979 13. NUMBER OF PAGES 96
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Approved for public release: IAW AFR 190-17

Joseph P. Hipps, Major, USAF
Director of Information

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Matrices (Mathematics)	Liapanov Equations	Generalized Inverses
Matrix Equations	MINQUE	Block Diagonalization
Linear Equations	Linear Algebraic Equations	
Reduced Order Filters	General Linear Equations	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Solutions of the general linear matrix equation $\sum_{i=1}^{n} A_i X B_i = C$ are obtained and presented in this thesis. Some special cases do arise like the Liapanov matrix equation. Necessary conditions and sufficient conditions are established for the solution of the general linear matrix equation. Other forms of solutions than

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

those obtained through the use of similarity transformations that have been considered make use of the spectral decomposition of matrices and tensor products of matrices or Kronecker products. In considering the general linear matrix equation, linear matrix equations in which two different variables appear are also studied. Conditions for the existence of a solution for this type of equation are given. The theory of the generalized inverses of a matrix was used in obtaining a solution to the general linear matrix equation. More general forms of the solution are given and conditions under which these solutions exist have been established. Solutions to systems of matrix equations were also considered. As a by-product of this investigation, some aspects of the model reduction problem may be treated from the point of view of matrix equations. In particular, a new method of solution of the matrix equation AXB + CXD = E which was recently considered by S.K. Mitra (Siam Journal of Applied Math, 32) and others was obtained. Applications of the results of this work are of use in the estimation of variance and covariance components of linear models as treated by C.R. Rao.

INSTRUCTIONS FOR PREPARATION OF REPORT DOCUMENTATION PAGE

RESPONSIBILITY. The controlling DoD office will be responsible for completion of the Report Documentation Page, DD Form 1473, in all technical reports prepared by or for DoD organizations.

CLASSIFICATION. Since this Report Documentation Page, DD Form 1473, is used in preparing announcements, bibliographies, and data banks, it should be unclassified if possible. If a classification is required, identify the classified items on the page by the appropriate symbol.

COMPLETION GUIDE

General. Make Blocks 1, 4, 5, 6, 7, 11, 13, 15, and 16 agree with the corresponding information on the report cover. Leave Blocks 2 and 3 blank.

Block 1. Report Number. Enter the unique alphanumeric report number shown on the cover.

Block 2. Government Accession No. Leave Blank. This space is for use by the Defense Documentation Center.

Block 3. Recipient's Catalog Number. Leave blank. This space is for the use of the report recipient to assist in future retrieval of the document.

Block 4. Title and Subtitle. Enter the title in all capital letters exactly as it appears on the publication. Titles should be unclassified whenever possible. Write out the English equivalent for Greek letters and mathematical symbols in the title (see "*Abstracting Scientific and Technical Reports of Defense-sponsored RDT/E*," AD-667 000). If the report has a subtitle, this subtitle should follow the main title, be separated by a comma or semicolon if appropriate, and be initially capitalized. If a publication has a title in a foreign language, translate the title into English and follow the English translation with the title in the original language. Make every effort to simplify the title before publication.

Block 5. Type of Report and Period Covered. Indicate here whether report is interim, final, etc., and, if applicable, inclusive dates of period covered, such as the life of a contract covered in a final contractor report.

Block 6. Performing Organization Report Number. Only numbers other than the official report number shown in Block 1, such as series numbers for in-house reports or a contractor/grantee number assigned by him, will be placed in this space. If no such numbers are used, leave this space blank.

Block 7. Author(s). Include corresponding information from the report cover. Give the name(s) of the author(s) in conventional order (for example, *John R. Doe* or, *if author prefers, J. Robert Doe*). In addition, list the affiliation of an author if it differs from that of the performing organization.

Block 8. Contract or Grant Number(s). For a contractor or grantee report, enter the complete contract or grant number(s) under which the work reported was accomplished. Leave blank in in-house reports.

Block 9. Performing Organization Name and Address. For in-house reports enter the name and address, including office symbol, of the performing activity. For contractor or grantee reports enter the name and address of the contractor or grantee who prepared the report and identify the appropriate corporate division, school, laboratory, etc., of the author. List city, state, and ZIP Code.

Block 10. Program Element, Project, Task Area, and Work Unit Numbers. Enter here the number code from the applicable Department of Defense form, such as the DD Form 1498, "Research and Technology Work Unit Summary" or the DD Form 1634, "Research and Development Planning Summary," which identifies the program element, project, task area, and work unit or equivalent under which the work was authorized.

Block 11. Controlling Office Name and Address. Enter the full, official name and address, including office symbol, of the controlling office. (*Equates to funding/sponsoring agency. For definition see DoD Directive 5200.20, "Distribution Statements on Technical Documents."*)

Block 12. Report Date. Enter here the day, month, and year or month and year as shown on the cover.

Block 13. Number of Pages. Enter the total number of pages.

Block 14. Monitoring Agency Name and Address (*if different from Controlling Office*). For use when the controlling or funding office does not directly administer a project, contract, or grant, but delegates the administrative responsibility to another organization.

Blocks 15 & 15a. Security Classification of the Report: Declassification/Downgrading Schedule of the Report. Enter in 15 the highest classification of the report. If appropriate, enter in 15a the declassification/downgrading schedule of the report, using the abbreviations for declassification/downgrading schedules listed in paragraph 4-207 of DoD 5200.1-R.

Block 16. Distribution Statement of the Report. Insert here the applicable distribution statement of the report from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 17. Distribution Statement (*of the abstract entered in Block 20, if different from the distribution statement of the report*). Insert here the applicable distribution statement of the abstract from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 18. Supplementary Notes. Enter information not included elsewhere but useful, such as: Prepared in cooperation with . . . Translation of (*or by*) . . . Presented at conference of . . . To be published in . . .

Block 19. Key Words. Select terms or short phrases that identify the principal subjects covered in the report, and are sufficiently specific and precise to be used as index entries for cataloging, conforming to standard terminology. The DoD "Thesaurus of Engineering and Scientific Terms" (*TEST*), AD-672 000, can be helpful.

Block 20. Abstract. The abstract should be a brief (*not to exceed 200 words*) factual summary of the most significant information contained in the report. If possible, the abstract of a classified report should be unclassified and the abstract to an unclassified report should consist of publicly- releasable information. If the report contains a significant bibliography or literature survey, mention it here. For information on preparing abstracts see "Abstracting Scientific and Technical Reports of Defense-Sponsored RDT&E," AD-667 000.

★ U.S. GOVERNMENT PRINTING OFFICE : 1973—729-091/143: 3-1