MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A138274

GENERALIZED INVERSES OF

MATRICES AND ITS APPLICATIONS

Thesis

/GCS/MA/ 　Abdel Monem E. Doma
AFIT/GA/GCS/83D-1 Lt Col Egyptian Army

DTIC
ELECTE
FEB 2 4 1984
S D
B

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

DTIC FILE COPY

Wright-Patterson Air Force Base, Ohio

84 02 21 163

GENERALIZED INVERSES OF

MATRICES AND ITS APPLICATIONS

Thesis

/GCS/MA/   Abdel Monem E. Doma
AFIT/MA/GCS/83D-1 Lt Col Egyptian Army

Approved for public release; distribution unlimited.

GENERALIZED INVERSES OF

MATRICES AND ITS APPLICATIONS

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

Abel-Monem E. Doma, B.S.

Lt Col   Egyptian Army

December 1983

Preface

My studies of optimization theory, control theory, and estimation problems prompted me to take a further look into improving solution techniques for matrix equations fundamental to those areas. This thesis is the result of this investigation using generalized inverses of that kind of matrices encountered in those areas.

It would not have been possible without the assistance and guidance of Dr. John Jones, Jr., whose imagination and fertile mind are the true source of this thesis.

This work was sponsored by the Aerospace Medical Research Laboratory (AMRL), Wright-Patterson AFB, Ohio. I would like to thank Dr. Dan Repperger of the AMRL for his support, and Major Castor for taking time to review this report. I also would like to thank Lt Col Edwards, whose notes and advice were a big help to complete this report.

Mrs. Becky Brumlow, my typist, deserves my most heartfelt thanks for taking this jumbled mass of equations and turning it into a clear, easily-readable format.

My wife, Amal and our sons, Ahmed and Tarek, have endured the most throughout this program of study. Their patience and love is without bound, and much appreciated.

Contents

Abstract

Theory and computation techniques of the various types of generalized inverses of matrices which have polynomial elements x, y, z..., etc., are presented. A simple algorithm for computation of generalized inverses of a constant matrix is established, and then applied to the case of matrices having polynomial elements in several variables. Reduction of a matrix to its Smith form over the ring of polynomial elements in several variables is presented. A simple algorithm for investigation of the system $Ax = b$ in case of constant and nonconstant rank of A is presented. Application of generalized inverses to solve more general matrix equations such as Lyapunov and Riccati equations is studied.

Introduction

The problem of solving m linear equations in n unknowns over the field of complex numbers can be formulated as a matrix equation $Ax = b$, where x is the column vector of 'unknowns'. If A is a non-singular square matrix, then the linear matrix equation has an immediate solution, given by $x = A^{-1}b$. If, however, A is a singular square matrix or in general nonsquare matrix, then the classical inverse of A is not defined. In this case it is possible to find a similar representation of the solutions of the system $Ax = b$ using generalized inverses associated with the matrix.

In 1920, Moore 22 published a theory for the generalized inverses in abstract form. In 1955, Penrose published a theory for a generalized inverse of any matrix with complex elements. He showed that for any matrix A whose elements are complex numbers there exists a unique generalized inverse, called Moore-Penrose inverse, A^+. This unique matrix A^+ is used to find the minimum-norm solution to the least-squares problem $||Ax-b|| = $ minimum. Penrose put four conditions which must be satisfied by the generalized inverse A^+.

Since 1955, the concept of generalized inverse has been modified to include more general generalized inverses which satisfy only some of Penroses's four

conditions. These generalized inverses are not unique. The most important generalized inverse, which is enough to investigate the system $Ax = b$, is an A_1 inverse which satisfies the first Penrose condition, $AA_1A = A$.

Chapter I of this thesis contains the theory of all types of generalized inverses. The theory of generalized inverse which produces a special solution for the system $Ax = b$, will be studied. In Chapter II, the computation techniques of the generalized inverses will be established. A simple algorithm for computing all types of generalized inverses will be introduced. In Chapter III, the theory of generalized inverses of matrices with polynomial elements will be discussed. The conditions under which a matrix with polynomial elements has a generalized inverse will be investigated. The problem of finding a solution of the system $Ax=b$, where A has polynomial elements in more than one variable, will be studied. Chapter IV will contain application of generalized inverse for solving more general matrix equations such as Lyapunov and Riccati equation.

Throughout this thesis capital letters denote matrices, while lower case letters represent scalars.

I Theory of Generalized Inverses of a Matrix

Definition of Generalized Inverses

The main purpose of this chapter is to establish the main concepts of generalized inverses of a matrix with complex elements. Any m·n matrix A having real or complex elements satisfies the following four axioms:

(I) $AXA = A$

(II) $XAX = X$

(III) (AX) is hermitian

(IV) XA is hermitian

where X is a unique (n·m) matrix. This unusual fact was proved by Moore$\begin{bmatrix} 22 \end{bmatrix}$ in 1920 in abstract form and established as shown above by R. Penrose$\begin{bmatrix} 25 \end{bmatrix}$ in 1955.

The introduction to this field of interest will be given in special cases with respect to the matrices which satisfy axiom I; axioms I, II; axioms I, II, III, and finally, axioms I, II, IV. These more general cases will be denoted by A_1, $A_{1,2}$, $A_{1,2,3}$, $A_{1,2,4}$, respectively, which in general, are not unique for given matrix A.

Below, is given a list of those types of generalized inverses. Unfortunately, there are a lot of different notations for each type, but in this thesis we will use the notation given above.

A_1-Inverses: every matrix possesses at least one A_1 inverse, sometimes termed "generalized inverse", A^- or g-inverse.

$A_{1,2,3,4}$-Inverse: Every matrix whose elements are real or complex numbers possesses a unique $A_{1,2,3,4}$ inverse called "Moore Penrose inverse" and usually denoted by A^+.

$A_{1,2}$-Inverse: Every matrix possesses at least one so-called weak generalized inverse, or reflexive generalized inverse.

$A_{1,2,3}, A_{1,2,4}$-Inverse: Every matrix whose elements are real or complex numbers possesses at least one $A_{1,2,3}$ and one $A_{1,2,4}$ inverse. In the case of full row rank (full column rank), there exists a right (left) inverse which in this case will be an $A_{1,2,3} (A_{1,2,4})$ inverse.

In this chapter we will try to investigate the theory of different kinds of generalized inverses. Let $C^{m \cdot n}$ denote the vector space of all m·n matrices having complex numbers as elements.

Characterization of A_1 Inverse

<u>Theorem (1-1)</u>: Let A be an m·n matrix whose elements are real or complex numbers (i.e., $A \epsilon C^{m \cdot n}$). The matrix G of order n·m is an A_1 generalized inverse of A if and only if $X = Gy$ is a solution for consistent equation $Ax = y$.

4

Proof: First, assume that the equation $Ax = y$ has a solution. This implies the existance of a vector w such that $Aw = y$. Substituting with $AA_1A=A$ we have

$$Aw=y \qquad AA_1(Aw)=y \qquad A(A_1y)=y$$

The last equation implies that $x=A_1y$ is a solution of $Ax=y$.

Second, suppose that $Ax = y$ is consistent. Suppose that there exists a solution X such that $x = Gy$ then,

$$A(Gy) = y$$

$$AGAx = y$$

$$AGA = A$$

G is A_1 inverse of matrix A.

Theorem (1-2): Let A be an $m \cdot n$ matrix whose elements are real or complex. Then, the matrices A_1A and AA_1 are idempotents with the same rank as A. Further, rank A = trace A_1A = trace AA_1 .

5

Proof: Suppose $H = A_1 A$ and $F = AA_1$

$$H^2 = A_1 (AA_1 A) = A_1 A = H$$

and also

$$F^2 = (AA_1 A)A_1 = AA_1 = F$$

Also, rank $A \geq$ rank $H \geq$ rank $AH =$ rank A and

rank $A \geq$ rank $F \geq$ rank $FA =$ rank A. These relationships

imply that rank $A =$ rank $H =$ rank F. Since H and F

are idempotent, the rank is given by the trace of each.

From the equation $AA_1 A = A_1$ it is clear that rank

$A_1 \geq$ rank A.

Based on previous theorem we can define A_1-inverse

of $A \epsilon C^{m \cdot n}$ as a matrix $A_1 \epsilon C^{n \cdot m}$ such that $(A_1 A)$ is an

idempotent and $R(A_1 A) = R(A)$ or, alternatively AA_1 is

idempotent and $R(AA_1) = R(A)$.

Characterization of $A_{1,2}$ Inverse

Theorem (1-3): The necessary and sufficient con-

dition for an A_1 inverse to be $A_{1,2}$ inverse of matrix

$A(m \cdot n)$ is that rank of $A_1 =$ rank of A.

Proof: Suppose that $A_{1,2}$ is an $\{1,2\}$-inverse of

A, then

$$AA_{1,2}A = A \qquad \text{rank } (A) \geq \text{rank } (A_{1,2})$$

$$A_{1,2}AA_{1,2} = A_{1,2} \qquad \text{rank } (A_{1,2}) \leq \text{rank } (A)$$

Thus, $\text{Rank}(A) = \text{rank } (A_{1,2})$

Conversely, suppose that $A_{1,2}$ is an $\{1\}$-inverse of A with $R(A) = R(A_{1,2})$. This implies the following

$$AA_{1,2}A = A$$

$$R(A) = R(A_{1,2}) = R(A_{1,2}A)$$

In addition, $A_{1,2}A$ is idempotent using the previous definition of $\{1\}$-inverse, A is $\{1\}$-inverse of $A_{1,2}$. This implies that $A_{1,2}AA_{1,2} = A_{1,2}$ and this completes the proof.

$\underline{\text{Theorem } (1\text{-}4)}$: Any $A_{1,2}$-inverse of A can be expressed as

$$A_{1,2} = \tilde{A}_1 A \tilde{\tilde{A}}_1$$

where $\tilde{A}_1$ and $\tilde{\tilde{A}}_1$ are (possibly different) A_1-inverses of A.

$\underline{\text{Proof}}$: The proof will be given in detail in Chapter II.

Matrices of Full Rank

Theorem (1-5): A matrix $A \epsilon C^{m \cdot n}$ has a right inverse A_r if and only if rank $(A) = m$ (full row rank). For a full row rank matrix $A \epsilon C^{m \cdot n}$, the following statements are equivalent:

a. A_r is a right inverse of A

b. A_r is an A_1-inverse of A

c. A_r is an $A_{1,2,3}$-inverse of A

Proof: To prove that "a" implies "b" and "c", suppose that A_r is a right inverse of A, i.e.,

$$AA_r = I_m$$

Premultiply previous relation by A_r and post-multiply by A we get

$$A_r AA_r = A_r, \text{ and}$$

$$AA_r A = A \text{ , respectively.}$$

Thus, A_r is an $A_{1,2,3}$-inverse. Next, suppose that A_r is a {1}-inverse of A. Since rank A equals m , we can find two nonsingular matrices P and Q such that

$$PAQ = \begin{bmatrix} I_m | 0 \end{bmatrix} \text{ , or}$$

$$A = P^{-1} \begin{bmatrix} I_m | 0 \end{bmatrix} Q^{-1}$$

any {1}-inverse can be written as follows:

$$A_1 = Q \left[\frac{I_m}{U}\right] P \quad,$$

where U is arbitrary.

Substituting in the first three axioms of generalized inverses, we get

$$AA_1A = (P^{-1}\left[I_m | 0\right]Q^{-1})Q\left[\frac{I_m}{U}\right]P(P^{-1}\left[I_m | 0\right]Q^{-1}$$

$$= P^{-1}\left[I_m | 0\right] = A$$

$$A_1AA_1 = (Q\left[\frac{I_m}{U}\right]P) \ P^{-1}\left[I_m | 0\right]Q^{-1}(Q \left[\frac{I_m}{U}\right]P)$$

$$= Q \left[\frac{I_m}{U}\right]P = A_1$$

$$AA_1 = P^{-1}\left[I_m | 0\right]Q^{-1} \ Q \left[\frac{I_m}{U}\right]P$$

$$= P^{-1} I_m P = I_m$$

Thus, "b" implies "a", and "c". It is clear also that "c" implies "a", and "b".

<u>Theorem (1-6)</u>: A matrix $A \epsilon C^{m \cdot n}$ has a left inverse (A_ℓ) if and only if, rank $(A) = n$. For a full column rank matrix $A \epsilon C^{m \cdot n}$, the following statements are equivalent:

a. A_ℓ is a left inverse of A

b. A_ℓ is an A_1-inverse of A.

c. A_ℓ is an $A_{1,2,4}$-inverse.

Proof: Proof is similar to theorem (1-5).

Minimum Norm and Least-squares Solution of Ax-y

Theorem (1-7): Let the norm of $x \in R^n$ be defined as $||x|| = (x*x)^{\frac{1}{2}}$. Let the equation Ax=y be consistent. Let the minimum norm solution x be x = Gy. Then, G is an $A_{1,2,4}$ inverse of A. In this case, the minimum norm solution of consistent equation Ax=y will be unique, although minimum norm generalized inverse may not be.

Proof: To illustrate the optimal property of $A_{1,2,4}$ as stated in the previous theorem, let us consider the following consistent system of equations

$$Ax = b$$

where

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \quad ; \quad b = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

One can check that the following matrices are $A_{1,2}, A_{1,2,4}, A_{1,2,4}$-inverses, respectively.

$$\begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & - \end{bmatrix} ; \begin{bmatrix} \dfrac{32}{109} & \dfrac{-11}{109} & 0 \\[2mm] \dfrac{96}{109} & \dfrac{-33}{109} & 0 \\[2mm] \dfrac{-75}{109} & \dfrac{36}{109} & 0 \\[2mm] \dfrac{7}{109} & \dfrac{1}{109} & 0 \end{bmatrix} , \begin{bmatrix} \dfrac{82}{109} & \dfrac{-31}{109} & \dfrac{10}{109} \\[2mm] \dfrac{246}{109} & \dfrac{-96}{109} & \dfrac{30}{109} \\[2mm] \dfrac{-90}{109} & \dfrac{42}{109} & \dfrac{-3}{109} \\[2mm] \dfrac{52}{109} & -17 & \dfrac{9}{109} \end{bmatrix} \quad .$$

Each of the above matrices will give a solution for the system as follows:

$$x_1 = \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \\ 0 \\ 3 \end{bmatrix}$$

$$x_2 = \frac{1}{109} \begin{bmatrix} 32 & -11 & 0 \\ 96 & -33 & 0 \\ -75 & 36 & 0 \\ 7 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = \frac{1}{109} \begin{bmatrix} -23 \\ -69 \\ 105 \\ 12 \end{bmatrix}$$

$$x_3 = \frac{1}{109} \begin{bmatrix} 82 & -31 & 10 \\ 246 & -96 & 30 \\ -90 & 42 & -3 \\ -52 & -17 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = \frac{1}{109} \begin{bmatrix} -23 \\ -69 \\ 105 \\ 12 \end{bmatrix} \quad ,$$

Notice that both $A_{1,2,4}$-inverses give the same solution.

Theorem (1-8): Let G be a matrix such that Gy is a least squares solution of the inconsistent equation $Ax = y$ for any $y \in R^m$. Then, G is an $A_{1,2,3}$-inverse of A. A least squares solution $x = Gy$ may not be unique, but minimum $||Ax-y||$ is unique. If Gy is a least squares solution, then the class of least squares solutions is $x_g = Gy + (I - GA)Z$, is arbitrary.

Proof: Note that the least squares solution is the solution of the consistent system $Ax = \tilde{y}$, where $\tilde{y}$ is the projection of y onto the column space of A. That means to find a solution to the system $Ax = AGy$ where G is an A_1-inverse substituting with general solution, we get

$$Ax = A(Gy + (I-G A)Z)$$

$$= AGy + AZ - AGAZ$$

$$= AGy + AZ - AZ$$

$$= AGy = \tilde{y} \ .$$

Notice that there are infinitely many least square solutions. Let us consider the following inconsistent linear system:

12

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A has the following $\{1,2,3\}$ generalized inverses.

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

These generalized inverses will give the following least square solutions

$$x_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} ; \quad x_2 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} ;$$

$$x_3 = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{matrix} 0 \\ 1 \end{matrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \end{bmatrix}$$

All have the same $||Ax - y|| = \frac{1}{\sqrt{2}}$

Theorem (1-.9): Let G be a matrix of order $n \cdot m$ such that Gy is the minimum-norm, least-squares solution of $Ax = y$. Then, G is the generalized inverse $A_{1,2,3,4} = A^{+}$ of the matrix A.

13

Proof: The proof is direct application on theorems (1-7) and (1-8). The previous example shows that the last solution (corresponding to $A_{1,2,3,4}$-inverse) is the minimum-norm, least-squares solution.

II Computation of Generalized Inverses

Introduction

The main purpose of this chapter is to establish
the techniques to be used in the more general cases of
computation of the various generalized inverses of
matrices.

First, the elements of matrices considered will be
real or complex numbers. Let $C^{m \cdot n}$ denote the vector space
of all $m \cdot n$ matrices having real or complex numbers as
elements. Let I_r be the identity matrix; that is,
$I_r \epsilon C^{r \cdot r}$. Let $0^{m \cdot n} \epsilon C^{m \cdot n}$ be the zero matrix, that is, all
elements equal zero. Capital letters will deonte matrices.

In the next section the general representations of
A_1 and $A_{1,2}$ are given. Reduction of a matrix $A \epsilon C^{m \cdot n}$
to its canonical form using elementary operations is used

to derive a general formula for computation of A_1 and $A_{1,2}$
generalized inverses.

In the following section, the techniques given above
will be modified. A simple computation technique for
computing A_1, $A_{1,2}$, $A_{1,2,3}$, $A_{1,2,4}$ and $A_{1,2,3,4}$ will
be given. This technique will be based on the previous
technique after suitable partitioning of the transformation
matrices.

Next, the basic technique given above will be applied
through simple examples.

The basic technique previously given will be applied to investigate the linear system of equations $Ax = b$ and to find the general solution, if it exists. A comparison between this new technique and the other techniques will be given through simple examples.

In the next section, the technique for computing A^+ using factorization will be given.

After that, the basic technique given in the third section will be applied for computing A_1 and A_2 of a specified rank.

General Representation of A_1 and $A_{1,2}$

In this technique, the Gausian elementary operations (row and column) will be used to reduce any matrix $A \epsilon C^{m \cdot n}$ to its canonical form. In other words, for each matrix $A \epsilon C^{m \cdot n}$ there exists two nonsingular matrices $R \epsilon C^{m \cdot m}$ and $C \epsilon C^{n \cdot n}$ such that

$$R \, A \, C = \bar{I} = \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} . \tag{2-1}$$

Theorem (2-1)

Let $A \epsilon C^{m \cdot n}$ with rank r. Let $R \epsilon C^{m \cdot m}$ and $C \epsilon C^{n \cdot n}$ be nonsingular such that

$$R \, A \, C = \begin{bmatrix} I_r & K \\ \hline 0 & 0 \end{bmatrix} \tag{2-2}$$

16

then

$$A_{1,2} = C \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R \qquad\qquad (2-3)$$

is a {1,2}inverse of A.

<u>Proof</u>:

First, we prove that $A_{1,2}$ satisfies the first axiom for generalized inverses as follows:

$$A\, A_{1,2}A = (R^{-1}\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1})\,(C\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R)\,(R^{-1}\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1}$$

$$= R^{-1}\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right]\, C^{-1}$$

$$= R^{-1}\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right]\, C^{-1}$$

$$= R^{-1}\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1} = A$$

Moreover, $A_{1,2}$ satisfies second axiom, that is

$$A_{1,2}A\, A_{1,2} = (C\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R)\,(R^{-1}\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] C^{-1})\,(C\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R)$$

$$= C\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]\, R$$

$$= C\left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R = C\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] R = A_{1,2}$$

17

Theorem (2-2)

Let $A \epsilon C^{m \cdot n}$ with rank r. Let R and C be as previously defined in (2-2). Then

$$A_1 = C \left[\begin{array}{c|c} I_r & U \\ \hline 0 & W \end{array}\right] R \qquad (2\cdot4)$$

where U and W are matrices of proper size. That is, $U \epsilon C^{r(m-r)}$ and $W \epsilon C^{(n-r)(m-r)}$.

Proof: The proof is trivial by substituting (2-4) into Axiom-1 of generalized inverse. That is:

$$A A_1 A = (R^{-1} \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1}) (C \left[\begin{array}{c|c} I_r & U \\ \hline V & W \end{array}\right] R) (R^{-1} \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1})$$

$$= R^{-1} \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} I_r & U \\ \hline 0 & W \end{array}\right] \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1}$$

$$= R^{-1} \left[\begin{array}{c|c} I_r & U+KW \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1} = R^{-1} \left[\begin{array}{c|c} I_r & K \\ \hline 0 & 0 \end{array}\right] C^{-1} = A$$

Theorem (2-3)

Let $A \epsilon C^{m \cdot n}$ and P, Q be nonsingular matrices such that:

$$PAQ = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \qquad (2-5)$$

where r is the rank of A. Then

$$A_1 = Q \left[\begin{array}{c|c} I_r & U \\ \hline V & W \end{array}\right] P; \qquad (2-6)$$

18

$$U \in C^{r \times (m-r)}, \quad V \in C^{(n-r) \times r}, \quad W \in C^{(n-r) \times (m-r)}$$

Proof: A_1 given in (2-6) satisfies the first axiom as follows:

$$A A_1 A = (P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1})(Q \left[\begin{array}{c|c} I_r & V \\ \hline V & W \end{array}\right] P)(P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1})$$

$$= P^{-1} \left[\begin{array}{c|c} I_r & U \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1}$$

$$= P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1} = A$$

Remark: There may be infinitely many A_1.

Theorem (2-4)

Let $A \in C^{m \cdot n}$ and let X belong to $C^{n \cdot m}$, then X is a {1,2}-inverse of A if and only if X has the following form:

$$X = \tilde{A}_1 A \tilde{\tilde{A}}_1 \qquad\qquad (2-7)$$

where $\tilde{A}_1$ and $\tilde{\tilde{A}}_1$ are two (not necessarily different) {1} -inverses for A.

Proof:

Sufficiency proof: if (2.7) holds, then X satisfies the first and second axioms. That is

$$AXA = A \ \tilde{A}_1 \ A\tilde{\tilde{A}}_1 A = A \ \tilde{A}_1 \ A = A$$

and

$$XAX = (\tilde{A}_1 \ A\tilde{\tilde{A}}_1)A \ (\tilde{A}_1 A\tilde{\tilde{A}}_1) = \tilde{A}_1 \ A\tilde{A}_1 A \ \tilde{A}_1$$

$$= \tilde{A}_1 \ A \ \tilde{\tilde{A}}_1 = A$$

Necessity proof: Let us suppose that X is a $\{1,2\}$ -inverse to A. X satisfies the second axiom. That is

$$X = XAX$$

using first axiom.

$$X = X(A \ X \ A)X \tag{2-8}$$
$$= X(AYA)X(AZA)X$$

where Y and Z are two $\{1\}$ -inverses for matrix A. X given in (2.8) can be reduced as follows:

$$X = XAY(AXA)ZAX$$
$$= (XAY) \ A \ (ZAX)$$
$$= N \ A \ M$$

To complete the proof, it is sufficient to prove that N and M are two $\{1\}$ -inverses for A as follows

$$ANA = A(XAY)A = (AXA)YA$$

$$= AYA = A$$

and

$$AMA = A(ZAX)A = (AZA)XA$$

$$= AXA = A.$$

Theorem (2-5)

Let $A_\varepsilon C^{m \cdot n}$ and let $X_\varepsilon C^{n \cdot m}$, then X is a {1,2}-inverse of A if and only if X has the following form:

$$X = Q \left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array} \right] P , \qquad (2-9)$$

where P,Q are two nonsingular matrices such that:

$$PAQ = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] \qquad (2-10)$$

and V,U are arbitrary matrices.

Proof: First, we prove that X given in (2.9) is {1,2}-inverse as follows:

$$AXA = (P^{-1}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1})\ (Q\left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array}\right] P)\ (P^{-1}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1})$$

$$= P^{-1}\left[\begin{array}{c|c} I_r & U \\ \hline O & O \end{array}\right]\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1}$$

$$= P^{-1}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1} = A$$

and

$$XAX = (Q\left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array}\right] P)\ (P^{-1}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1})\ (A\left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array}\right] P)$$

$$= Q\left[\begin{array}{c|c} I_r & O \\ \hline V & O \end{array}\right]\left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array}\right] P$$

$$= Q\left[\begin{array}{c|c} I_r & U \\ \hline V & VU \end{array}\right] P = X$$

Second, suppose that X is {1,2}-inverse, then, by using theorem (2-4) X can be expressed as:

$$X = \tilde{A}_1\ A\ \tilde{A}_2 = (Q\left[\begin{array}{c|c} I_r & U' \\ \hline V' & W' \end{array}\right] P)\ (P^{-1}\left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right] Q^{-1})\ (Q\left[\begin{array}{c|c} I_r & U'' \\ \hline V'' & W'' \end{array}\right] P)$$

$$= Q\left[\begin{array}{c|c} I_r & O \\ \hline V' & O \end{array}\right]\left[\begin{array}{c|c} I_r & U'' \\ \hline V'' & W'' \end{array}\right] P$$

$$= Q\left[\begin{array}{c|c} I_r & U'' \\ \hline V' & V'U'' \end{array}\right] P\ \ .$$

That means that X has the form given in (2-9).

Lemma (2-1)

Let $X \epsilon C^{n \cdot m}$ and $A \epsilon C^{m \cdot n}$ and let X be

$$X = Q \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] P \qquad (2\text{-}11)$$

then X is a $\{1,2\}$-inverse of A .

Proof: Proof is evident by choosing $U = 0$, and $V = 0$ in (2-9)

The Basic Technique

In this part, the basic algorithm used for computation of generalized inverses will be established. Let $C^{m \cdot n}$ denote the vector space of all $m \cdot n$ matrices having real or complex elements. Let I_r be the identity matrix; that is, $I_r \epsilon C^{r \cdot r}$. Let $0^{m \cdot n} \epsilon C^{m \cdot n}$ be the zero matrix; that is, all elements equal zero. Capital letters will denote matrices.

Theorem (2-6): Let $A \epsilon C^{m \cdot n}$ with rank r. Let $P \epsilon C^{m \cdot m}$, $Q \epsilon C^{n \cdot n}$ be nonsingular matrices such that

$$PAQ = \bar{I} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] . \qquad (2\text{-}12)$$

Consider the following partitioning for matrices P and Q

23

$$P = \left[\frac{T}{M}\right] \quad ; \quad Q = \left[S \mid N\right] \qquad (2\text{-}13)$$

where $T \epsilon C^{r \cdot m}$; $M \epsilon C^{(m-r) \cdot m}$; $S \epsilon C^{r \cdot r}$, $N \epsilon C^{n(n-r)}$.

Then the following pair of matrices

$$\left[\begin{array}{c|c} A & I_m \\ \hline I_n & 0 \end{array}\right] \quad ; \quad \left[\begin{array}{c|c|c} I_r & 0 & T \\ \hline 0 & 0 & M \\ \hline S & N & 0 \end{array}\right] \qquad (2\text{-}14)$$

are equivalent.

Proof: For any $A \epsilon C^{m \cdot n}$ there exists nonsingular matrices P, Q as defined in (2-12). To show that the matrices given in (2-14) are equivalent:

$$\left[\begin{array}{c|c} P & 0 \\ \hline 0 & I_n \end{array}\right] \left[\begin{array}{c|c} A & I_m \\ \hline I_n & 0 \end{array}\right] \left[\begin{array}{c|c} Q & 0 \\ \hline 0 & I_m \end{array}\right] = \left[\begin{array}{c|c} PA & P \\ \hline I_n & 0 \end{array}\right] \left[\begin{array}{c|c} Q & 0 \\ \hline 0 & I_m \end{array}\right]$$

$$= \left[\begin{array}{c|c} PAQ & P \\ \hline Q & 0 \end{array}\right] = \left[\begin{array}{c|c|c} I_r & 0 & T \\ \hline 0 & 0 & M \\ \hline S & N & 0 \end{array}\right]$$

Theorem (2-7): Let $A \epsilon C^{m \cdot n}$, let P, Q, S, T, M, N be matrices defined as in (2-12) and (2-13), then $A_{1,2} = (ST)$ is a $\{1,2\}$-inverse of A.

24

Proof: Applying Lemma (2-1)

$$A_{1,2} = Q \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] P$$

$$= \left[S \mid N \right] \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] \left[\frac{T}{M} \right] \quad .$$

$$= \left[S \mid N \right] \left[\frac{T}{0} \right] = ST$$

Theorem (2-8): Let the matrices A, P, Q, S, T, M, N be matrices defined as in theorem (2-7). Moreover, let

$$TM^t = 0, \tag{2-15}$$

then

$$A_{1,2,3} = ST$$

Proof: First, let $X = ST$ and let $TM^t = 0$. It is obvious that X satisfies the first and second axioms. To complete the proof it is sufficient to verify the third axiom. That is

$$AX = (P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] Q^{-1})(ST) \tag{2-16}$$

$$= P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] (Q^{-1}S)T$$

Using the fact that $Q^{-1}Q = I_n$, then

$$Q^{-1} \begin{bmatrix} S & | & N \end{bmatrix} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & I_{n-r} \end{array}\right]$$

$$\begin{bmatrix} Q^{-1}S & | & Q^{-1}N \end{bmatrix} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & I_{n-r} \end{array}\right]$$

$$Q^{-1}S = \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] ; \qquad Q^{-1}N = \left[\begin{array}{c} 0 \\ \hline I_{n-r} \end{array}\right] \qquad (2-17)$$

Substituting (2-17) into (2-16) we can write

$$AX = P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] \quad T \qquad (2-18)$$

$$= P^{-1} \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right] \quad T \quad = \quad P^{-1} \left[\begin{array}{c} T \\ \hline 0 \end{array}\right]$$

But

$$P \, P^t = \left[\begin{array}{c} T \\ \hline M \end{array}\right] \begin{bmatrix} T^t & | & M^t \end{bmatrix} = \left[\begin{array}{c|c} TT^t & TM^t \\ \hline MT^t & MM^t \end{array}\right] \qquad (2-19)$$

Substituting with (2-15) into (2-19)

26

$$P P^t = \left[\begin{array}{c|c} TT^t & 0 \\ \hline 0 & MM^t \end{array} \right]$$

multiplying both sides by

$$\left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & (MM^t)^{-1} \end{array} \right]$$

we have

$$P(P^t \left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & (MM^t)^{-1} \end{array} \right]) = I_m$$

i.e.,

$$P^{-1} = P^t \left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & (MM^t)^{-1} \end{array} \right]$$

Substituting with P^{-1} into (2-18) we have

$$AX = P^t \left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & (MM^t)^{-1} \end{array} \right] \left[\begin{array}{c} \left[\begin{array}{c|c} I_r & 0 \end{array} \right] P \\ \hline 0 \end{array} \right]$$

$$= P^t \left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & (MM^t)^{-1} \end{array} \right] \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] P$$

$$= P^t \left[\begin{array}{c|c} (TT^t)^{-1} & 0 \\ \hline 0 & 0 \end{array} \right] P$$

which is symmetric.

Theorem (2-9): Let the matrices A,P,Q,S,T,M,N
be matrices defined as in theorem (2-6). Moreover, let

$$N^t S = 0 ; \qquad \qquad (2-21)$$

then $A_{1,2,4} = ST$.

Proof: The proof is similar to theorem (2-8).

$$A_{1.2.4}A = (ST)\ P^{-1} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1} \qquad (2-22)$$

$$= S\ \left[\begin{array}{c|c} I_r & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1}\ , \text{ where } TP^{-1} = \left[\begin{array}{c|c} I_r & 0 \end{array}\right]$$

$$= Q\ \left[\begin{array}{c} I_r \\ \hline 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \end{array}\right]\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] Q^{-1}$$

$$Q^t Q = \left[\begin{array}{c} S^t \\ \hline N^t \end{array}\right]\ \left[\begin{array}{c|c} S & N \end{array}\right] = \left[\begin{array}{c|c} S^t S & S^t N \\ \hline N^t S & N^t N \end{array}\right] = \left[\begin{array}{c|c} (S^t S) & 0 \\ \hline 0 & (N^t N) \end{array}\right]$$

i.e.,

$$Q^{-1} = \left[\begin{array}{c|c} (S^t S)^{-1} & 0 \\ \hline 0 & (N^t N)^{-1} \end{array}\right]\ Q^t$$

Substituting into (2-21) we have

$$A_{1,2,4}A = Q\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} (S^t S)^{-1} & 0 \\ \hline 0 & (N^t N)^{-1} \end{array}\right] \qquad (2-23)$$

$$= Q\ \left[\begin{array}{c|c} (S^t S)^{-1} & 0 \\ \hline 0 & 0 \end{array}\right]\ Q^t$$

28

which is symmetric.

Lemma (2-1): Let the matrices A,P,Q,S,T,M,N be matrices as in theorem (2-7). Moreover, let

$$T M^t = 0$$

and

$$N^t S = 0 \quad \text{hold},$$

then

$$A^t = A_{1,2,3,4} = ST$$

Proof: Referring to theorems (2-8) and (2-9), it is clear that A^+ satisfies the four axioms.

Application of the Basic Technique

The following examples will illustrate the computation techniques given in the second section. This technique will be generalized to the cases where the elements of the matrix are polynomials in several parameters.

Example 1. Given the following matrix .

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2-24}$$

It is required to compute $A_1, A_{1,2}, A_{1,2,3}, A_{1,2,3,4}$ (A^+). The first step in the process is to form the following array

29

$$\left[\begin{array}{c|c} A & I_m \\ \hline I_n & 0 \end{array}\right] = \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] \tag{2-25}$$

Elementary row operations on A will affect I_m only and elementary column operations on A will affect the columns of I_n only. A obviously does not have a classical inverse, since it is not square. The second step is to perform row elementary operations on A to get the following array

$$\left[\begin{array}{ccc|cc} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] \tag{2-26}$$

Next step is to perform elementary operations on the columns of A to get the following array

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ \hline 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] \triangleq \left[\begin{array}{cc|c} I_r & 0 & P \\ 0 & 0 & \\ \hline Q & & 0 \end{array}\right] \tag{2-27}$$

30

$$\triangleq \begin{bmatrix} I_r & 0 & T \\ \hline 0 & 0 & M \\ \hline S & N & 0 \end{bmatrix} \qquad (2\text{-}28)$$

This process is not unique, however, one should check the following for possible errors

$$PAQ = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{bmatrix}$$

Now we can compute $A_{1,2}$ as follows

$$A_{1,2} = ST = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

We can check the answer to see that

$$A\, A_{1,2}\, A = A\ ,$$

and

$$A_{1,2}\, A\, A_{1,2} = A_{1,2}$$

Next, to get $A_{1,2,3}$ it is required to further modify (2-27) array by making the rows of T orthogonal to rows of M. This can be accomplished by adding multiples of second row of A to the first row. In our example we will

31

add ($\frac{1}{2}$) times the second row to the first one and this will result

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & 0 & 0 & -1 & 1 \\ \hline 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right]$$

Then

$$A_{1,2,3} = ST = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

To check the last result

$$AA_{1,2,3} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

which is symmetric.

Next, when the first column of the array is made to be orthogonal to the second and third one by adding 1/3 times second column and 1/3 times third colum to the first, we have the following array

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

Now

$$
A_{1,2,3,4} = A^{+} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} \end{bmatrix} = \begin{bmatrix} 1/6 & 1/6 \\ 1/6 & 1/6 \\ 1/6 & 1/6 \end{bmatrix}
$$

Example 2: Let

$$
A = \begin{bmatrix}
10 & 14 & 28 & 32 \\
8 & 8 & 4 & 8 \\
4 & 12 & 48 & 48
\end{bmatrix}
$$

After performing column and row elementary operations, the array will be as follows

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1/4 & 0 \\
0 & 1 & 0 & 0 & \tfrac{1}{2} & \tfrac{1}{2} & 0 \\
0 & 0 & 0 & 0 & -2 & 2 & 1 \\
0 & 1 & 12 & -3 & & & \\
0 & 0 & 0 & 1 & & & \\
1 & 2 & 22 & -4 & & 0 & \\
0 & -2 & -23 & 4 & & &
\end{bmatrix}
\triangleq
\begin{bmatrix}
1 & 0 & 0 & 0 & & & \\
0 & 1 & 0 & 0 & & & T \\
0 & 0 & 0 & 0 & & & M \\
 & & & & & & \\
 & & & & & & O \\
S & N & & & & &
\end{bmatrix}
$$

Making the rows of T orthogonal to that of M by adding multiple of third row to rows of T we have

$$
\left[\begin{array}{cc|cc||ccc}
1 & 0 & 0 & 0 & \frac{1}{9} & \frac{5}{36} & \frac{-1}{18} \\
0 & 1 & 0 & 0 & \frac{-23}{18} & \frac{-31}{18} & \frac{+16}{18} \\
\hline
0 & 0 & 0 & 0 & -2 & 2 & 1 \\
\hline
0 & 1 & 12 & -3 & & & \\
0 & 0 & 0 & 1 & & & \\
1 & 2 & 22 & -4 & & & \\
0 & -2 & -23 & 4 & & &
\end{array}\right]
\triangleq
\left[\begin{array}{cc|cc||c}
1 & 0 & 0 & 0 & \tilde{T} \\
0 & 1 & 0 & 0 & \\
\hline
0 & 0 & 0 & 0 & M \\
\hline
 & & & & \\
S & & N & & 0 \\
 & & & &
\end{array}\right]
$$

Finally, we can make the column of S to be orthogonal to that N (i.e., $S^{t}N=0$). Using multiples of columns of N, we get the following array

$$
\left[\begin{array}{cc|cc||ccc}
1 & 0 & 0 & 0 & \frac{1}{9} & \frac{5}{36} & \frac{-1}{18} \\
0 & 1 & 0 & 0 & \frac{-23}{18} & \frac{-31}{18} & \frac{16}{18} \\
\hline
0 & 0 & 0 & 0 & -2 & 2 & 1 \\
\hline
 & & 12 & -3 & & & \\
\tilde{S} & & 0 & 1 & & & \\
 & & 22 & -4 & & & \\
 & & -23 & 4 & & &
\end{array}\right]
\triangleq
\left[\begin{array}{cc|cc||c}
1 & 0 & 0 & 0 & T \\
0 & 1 & 0 & 0 & \\
\hline
0 & 0 & 0 & 0 & M \\
\hline
 & & & & \\
S & & N & & 0 \\
 & & & &
\end{array}\right]
$$

Now, we can calculate all kinds of generalized inverses.

To find $A_{1,2,3}$-inverse or $A_{1,2,4}$ inverse using the technique given in theorem (2-8) and (2-9), we have to establish a technique to find T or S such that (2-15) or (2-21) holds. That means to find $A_{1,2,3}$-inverse we have to find $\tilde{T}$ such that

$$\tilde{T} M^t = 0 \qquad\qquad (2\text{-}30)$$

The direct way to find T is to add a multiple of rows of M matrix such that (2-30) holds because this will not affect the canonical form given in (2-14). In matrix form we can write T as

$$\tilde{T} = T + K M$$

where K is chosen such that (2-30) holds, i.e.,

$$(T + K M) M^t = 0 .$$

This implies that

$$K = -TM^t (MM^t)^{-1}$$

Now the following theorem is clear.

Theorem (2-10): Let $A \in C^{m \cdot n}$, M,T,S,N be defined as in (2-14). Then the following holds

a) $A_{1,2} = S\ T$

b) $A_{1,2,3} = S\ \tilde{T} = S(T-TM^t(MM^t)^{-1}M)$

c) $A_{1,2,4} = \tilde{S}\ T = (S - N(N^tN)^{-1}N^tS)T$

d) $A_{1,2,3,4} = \tilde{S}\ \tilde{T}$

<u>Proof:</u>

It is sufficient to prove only part C. From theorem(2-9) it is sufficient to calculate S such that

$$\tilde{S}^{T}N = 0 \qquad\qquad (2-31)$$

holds. This can be done by adding multiples of columns of N to that of S. That is,

$$\tilde{S} = S + NK \qquad\qquad (2-32)$$

Substituting (2-31) we have

$$(S + NK)^t\ N = 0$$
$$(S^tN + K^tN^tN) = 0$$
$$K^t(N^tN) = -S^tN$$

But since (N^tN) is nonsingular matrix, we have

$$K^t = -S^tN(N^tN)^{-1}$$
$$K = -(N^+N)^{-1}N^tS$$

Substituting into (2-32) we have

$$\tilde{S} = S + NK$$
$$= S - N(N^tN)^{-1}N^tS$$

that is $A_{1,2,4}$ will be

$$A_{1,2,3} = \tilde{S}\,T$$
$$= (S - N(N^tN)^{-1}N^tS)T$$
$$= ST - N(N^tN)^{-1}N^tST$$
$$= (S - N(N^tN)^{-1}N^tS)T.$$

To apply this theorem let us consider the example ( ). T can be calculated as follows

$$T = T - TM^t(M\,M^t)^{-1}M$$

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

$$= \begin{vmatrix} 1 & 0 \end{vmatrix} + \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 2 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 \end{bmatrix}$$

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

In the same way $\tilde{S}$ can be calculated as follows

$$\tilde{S} = S - N(N^tN)^{-1} N^tS$$

$$N^tN = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$(N^tN)^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

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

$$= \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

To apply that to another example, let us consider example (2) to calculate $\tilde{S}$.

$$(N^tN)^{-1} = \begin{bmatrix} 12 & 0 & 22 & -23 \\ -3 & 1 & -4 & 4 \end{bmatrix} \begin{bmatrix} 12 & -3 \\ 0 & 1 \\ 22 & -4 \\ -23 & 4 \end{bmatrix} = \begin{bmatrix} 1157 & -216 \\ -216 & 42 \end{bmatrix}$$

$$(N^tN)^{-1} = \frac{1}{2(969)} \begin{bmatrix} 42 & +216 \\ +216 & 1157 \end{bmatrix}$$

$$\tilde{S} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 2 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 12 & -3 \\ 0 & 1 \\ 22 & -4 \\ -23 & 4 \end{bmatrix} \frac{1}{2(969)} \begin{bmatrix} 42 & 216 \\ 216 & 1157 \end{bmatrix} \begin{bmatrix} 12 & 0 & 22 & -23 \\ -3 & 1 & -4 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 2 \\ 0 & -2 \end{bmatrix}$$

$$= \frac{1}{(969)} \begin{bmatrix} -174 & 4797 \\ -62 & -6388 \\ 557 & -626 \\ 442 & 1904 \end{bmatrix}$$

The General Solution of the System

$\underline{Ax = b}$: Let $A \epsilon C^{m \cdot n}$ and $b \epsilon C^{m \cdot 1}$. Let P, Q, M, N, T, S be defined as in Section 2.2. It is proved in Appendix A theorem(A-1) that there exists a solution for the system $Ax = b$ if and only if

$$A A_1 b = b .$$

In this case the general solution is

$$X = A_1 b + (I_n - A_1 A) Z ,$$

where Z is arbitrary. Using the special generalized inverse given in previous section, i.e. $A_1 = ST$, we can check the consistency condition of the system $Ax = b$ and moreover, we can find the general solution in an easier way. This is clear from the following theorem.

<u>Theorem (2-11)</u>: The equation $Ax=b$ for $A \in C^{m \cdot n}$ and $b \in C^{m \cdot 1}$, has a solution $X \in C^{n \cdot 1}$ if and only if

$$Mb = 0 \qquad\qquad (2-33)$$

In this case the general solution is given by

$$X = (ST)b + NZ \qquad\qquad (2-34)$$

where the matrix Z is arbitrary and the matrices M,S,T,N are defined as in Section (2-2).

<u>Proof:</u> For any $A \in C^{m \cdot n}$ there exists nonsingular matrices P and Q such that the following holds

$$PAQ = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right]$$

$Ax = b$ has a solution x iff

$PAX = Pb$ has a solution x iff

$PAQQ^{-1}x = Pb$ has a solution x iff

$(PAQ)y = Pb$ has a solution y; $x=Qy$ iff

$$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c} W \\ \hline Z \end{array} \right] = \left[\begin{array}{c} T \\ \hline M \end{array} \right] b \quad ; \qquad P = \left[\begin{array}{c} T \\ \hline M \end{array} \right] \quad ; \quad y = \left[\begin{array}{c} W \\ \hline Z \end{array} \right]$$

$$x = Qy = \left[\begin{array}{c|c} S & N \end{array} \right] \left[\begin{array}{c} W \\ \hline Z \end{array} \right] = SW + NZ \qquad \text{iff}$$

40

$$\left[\frac{W}{0}\right] = \left[\frac{Tb}{Mb}\right] \quad ; \quad X = SW + NZ \qquad iff$$

$Mb = 0$; $X = STb + NZ$ where Z is an arbitrary matrix of appropriate size.

Remark: The solution $X = (ST)b$ for the system $Ax = b$ is a particular solution of $Ax = b$, and NZ is the general solution of the homogenious equation $Ax = 0$.

Example (3): Consider the system $Ax = b$ with

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

To find the general solution using both methods, we must calculate the generalized inverse first.

Using the elementary operation S, we can write the following array:

$$\begin{bmatrix} \begin{array}{cc|cc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 5 & -2 & 1 \\ \hline 1 & -2 & 3 & -3 & & & \\ 0 & 0 & 0 & 1 & & 0 & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 1 & -3 & 0 & & & \end{array} \end{bmatrix} = \begin{bmatrix} \begin{array}{c|c|c} I_2 & 0 & T \\ \hline 0 & 0 & M \\ \hline S & N & 0 \end{array} \end{bmatrix}$$

To check the first step

$$PAQ = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 3 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 8 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_{1,2} = ST = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix}$$

To check the consistency condition

$$A \, A_1 b = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

Thus, the system is consistent and has the following
general solution:

$$X = A_1 b + (I_n - A_1 A)Z$$

$$= \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} + \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \right) Z$$

$$= \begin{bmatrix} -5 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 & -3 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -3 & 0 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ Y \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} V + \begin{bmatrix} 3 \\ 0 \\ 1 \\ -3 \end{bmatrix} W$$

Using the second method, we can first check the consistency condition as follows:

$$Mb = \begin{bmatrix} 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = 0 \quad.$$

The general solution is

$$X = (ST)b + NZ$$

$$= \begin{bmatrix} 5 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 & -3 \\ 0 & 1 \\ 1 & 0 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} V \\ W \end{bmatrix}$$

$$
= \begin{bmatrix} -5 \\ 0 \\ 0 \\ 3 \end{bmatrix} \quad +V \begin{bmatrix} 3 \\ 0 \\ 1 \\ -3 \end{bmatrix} \quad +W \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}
$$

It is easy to see from previous calculation that the second method is much simpler.

<u>Remark</u>: The columns of N matrix are the basis for the null space of A.

Another Technique for Computation of A^+

<u>Theorem (2-12)</u>: Let $A \epsilon C^{m \cdot n}$ and have rank $k \leq$ min (m,n). Let $A = BC$, where rank of B = rank of C = k ; $B \epsilon C^{m \cdot n}$ and $C \epsilon C^{r \cdot n}$. Then $(B^t B)^{-1}$, and $(C \ C^t)^{-1}$ exist and

$$
A_{1,2,3,4} = A^+ = C^t (C \ C^t)^{-1} (B^+ B)^{-1} B^t \tag{2-35}
$$

<u>Proof:</u> Substituting in the four axioms by (2-35), we have

$$
AA^+ A = B(C \ C^t)(C \ C^t)^{-1} (B^t B)^{-1} (B^t B)^{-1} (B^t B) C
$$

$$
= BC = A
$$

and

44

$$A^+AA^+ = C^t(CC^t)^{-1}(B^tB)^{-1}(B^tB)(C\ C^t)(CC^t)^{-1}(B^tB)^{-1}B^t$$

$$= C^t(CC^t)^{-1}(B^tB)^{-1}B^t = A^+$$

and

$$AA^+ = BC\ C^t(CC^t)^{-1}(B^tB)^{-1}B^t$$

$$= B(B^tB)^{-1}B^t \text{ is symmetric}$$

and finally

$$A^+A = C^t(CC^t)^{-1}(B^tB)^{-1}B^+B\ C$$

$$= C^t(CC^t)^{-1}C$$

Theorem (2-13): Let $A \in C^{m \cdot n}$, let P, Q be nonsingular matrices such that

$$PAQ = \bar{I} = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \quad , \quad \text{then}$$

$$A = BC \quad , \quad B \in C^{m \cdot r}, \quad C \in C^{r \cdot n}_r$$

where
$$B = P^{-1}\left[\frac{I_r}{0}\right]$$

$$C = \left[\,I_r\,\middle|\,0\,\right]\quad Q^{-1}\ .$$

Proof:

$$BC = P^{-1} \begin{bmatrix} I_r \\ \hline 0 \end{bmatrix} \begin{bmatrix} I_r & | & 0 \end{bmatrix} Q^{-1}$$

$$= P^{-1} \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1} = A$$

Generalized Inverses of a Specified Rank

The purpose of this section is to compute A_1 and A_2 of a specified rank. In chapter I we proved that the rank of A_1 must be greater than or equal to the rank of the matrix, i.e.,

$$r \leq R(A_1) \leq \min(m,n).$$

It is clear that rank of A_2 will be less than or equal rank of A.

Now let $A \epsilon C^{m \cdot n}$. Let P and Q be two nonsingular matrices such that

$$PAQ = \begin{bmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{bmatrix}.$$

Let P and Q be partitioned as follows

$$P = \begin{bmatrix} T^{(K)} \\ \hline M^{(K)} \end{bmatrix} , \quad Q = \begin{bmatrix} S^{(K)} & | & N^{(K)} \end{bmatrix} \qquad (2\text{-}36)$$

46

where $T^{(k)} \epsilon C^{k \cdot m}$, $M^{(k)} \epsilon C^{(m-k)m}$ $S^{(k)} \epsilon C^{n \cdot k}$ and $M^{(k)} \epsilon C^{n(n-k)}$ where $0 \leq k \leq \min(m,n)$ and the following theorem uses the partitioning given in (2-36) to compute A_1 and A_2 of specified rank k.

Theorem (2-13): Let $A \epsilon C^{m \cdot n}$. Let P and Q be non-singular matrices such that

$$PAQ = \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} .$$

Let $T^{(k)}$, $M^{(k)}$, $S^{(k)}$, and $N^{(k)}$ be given as in (   ). Let $A_1^{(k)}$ and $A_2^{(k)}$ denote A_1 and A_2 of specified rank k, then

$$A_1^{(k)} = S^{(k)} T^{(k)} \qquad r \leq k \leq \min(m,n)$$

$$A_2^{(k)} = S^{(k)} T^{(k)} \qquad u \leq k \leq r$$

$$A_{1,2} = S^{(r)} T^{(r)} = ST$$

where S and T are given as in Section (2-3).

Proof: The proof will be easy if we notice that

$$S^{(k)} T^{(k)} = Q \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} P \qquad (2-37)$$

To complete the proof, substitute (2-37) into the first and second conditions of generalized inverse as follows:

$$A \, A_1^{(k)} \, A = P^{-1} \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1} Q \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} P P^{-1} \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1}$$

$$= \begin{cases} P^{-1} \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1} = A \quad \text{for} \quad k \geq r \\[3em] P^{-1} \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1} \quad\quad \text{for} \quad k \leq r \end{cases}$$

This implies that

$$A_1^{(k)} = S^{(k)} T^{(k)} ; \quad k \geq r$$

Substituting in the second condition we have

$$A_2^{(k)} \, A \, A_2^{(k)} = Q \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} P P^{-1} \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} Q^{-1} Q \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} P$$

$$= \begin{cases} Q \begin{bmatrix} I_k & 0 \\ \hline 0 & 0 \end{bmatrix} P = A_2^{(k)} \quad\quad \text{for } k \leq r \\[3em] Q \begin{bmatrix} I_r & 0 \\ \hline 0 & 0 \end{bmatrix} P \quad\quad\quad \text{for } k \geq r \end{cases}$$

This implies that

$$A_2^{(k)} = S^{(k)} T^{(k)} ; \quad k \leq r$$

Example (4):

Given

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 2 & 6 \\ 3 & 3 & 4 \end{bmatrix}$$

then, one can construct the following array

$$\begin{bmatrix} 1 & 0 & 0 & \vline & 1 & 0 & 0 \\ 0 & 1 & 0 & \vline & 2/3 & -1/6 & 0 \\ 0 & 0 & 0 & \vline & -1 & -1/2 & 1 \\ 1 & -2 & -5/3 & \vline & & & \\ 0 & 1 & 1/3 & \vline & & & \\ 0 & 0 & 1 & \vline & & & \end{bmatrix}$$

From this array we can write

$$S^{(0)} \triangleq 0 ,$$

$$S^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} ,$$

$$S^{(2)} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} ,$$

49

$$S^{(3)} = \begin{bmatrix} 1 & -2 & -5/3 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{bmatrix} \quad ,$$

$$T^{(0)} = 0 \quad ,$$

$$T^{(1)} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \quad ,$$

$$T^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & -1/6 & 0 \end{bmatrix}$$

and finally

$$T^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & -1/6 & 0 \\ -1 & -1/2 & 1 \end{bmatrix}$$

Now the generalized inverses can be computed as follows:

$A_2^{(0)} = 0$, the trivial solution for XAX=X.

$$A_2^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad ,$$

$$A_{1,2} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & -1/6 & 0 \end{bmatrix} \qquad r = 2$$

$$= \begin{bmatrix} 1/3 & 1/3 & 0 \\ 2/3 & -1/6 & 0 \\ 0 & 0 & 0 \end{bmatrix} ,$$

$$A_1^{(3)} = \begin{bmatrix} 1 & -2 & -5/3 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & -1/6 & 0 \\ -1 & -1/2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4/3 & 7/6 & -5/3 \\ 1/3 & -1/3 & 1/3 \\ -1 & -1/2 & 1 \end{bmatrix}$$

III <u>Matrices Over a Polynomial Ring</u>

<u>in Several Variables</u>

Introduction

The purpose of this chapter is to establish an algorithm to compute the generalized inverses for the matrices whose elements belong to a polynomial ring in several variables having coefficients which are complex numbers. In the previous chapter, we established many algorithms to compute the generalized inverses for constant matrices using reduction of the matrix to the following canonical form:

$$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right]$$

(3-1)

In the case of matrices with polynomial elements in several variables, it is not true that all matrices can be reduced to the form (3-1). To study the existence of generalized inverses of those matrices, we have to reduce the matrix to another simple form called the Smith Normal form. Thus, the study of the existence of the Smith form for matrices over the polynomial ring will be a necessary step to characterize those matrices which admit generalized inverses and to compute them in case of their existence.

In the next section, we will study the existence of the Smith form for matrices with polynomial elements. Conditions under which a matrix over a polynomial ring is

equivalent to its Smith form will be investigated. A systematic algorithm to reduce a matrix to its Smith form will be established.

In the third section, the necessary and sufficient conditions for a matrix over polynomial ring to have a certain type of generalized inverse will be given. The computation of generalized inverses for matrices with constant rank size having Smith form as in (3-1), will be established.

In the fourth section, the case of variable rank matrices will be investigated. The problem of finding solutions of the system $Ax = b$ in the case of variable rank matrices will be discussed, since it was not treated before in any preceeding work.

Throughout this chapter, C will denote the field of complex numbers. Let $R = C\left[\theta_1, \theta_2, \ldots, \theta_n\right]$ be the ring of polynomials in the variables $\theta_1, \theta_2, \ldots, \theta_n$ with coefficients belonging to C. Let $C^{m \cdot n}(\theta)$ be the vector space of all matrices of order $m \cdot n$ and their elements belong to R.

Reduction to Smith Form

Consider any $m \cdot n$ matrix $A(\theta) \epsilon\ C^{m \cdot n}(\theta)$. The Smith form $S(\theta) \epsilon\ C^{m \cdot n}(\theta)$ of the matrix $A(\theta)$ is defined to be

$$S(\theta) = \left[\text{diag}\ \left[e_i(\theta) \right] \middle|\ 0 \right] \quad ; \quad m < n$$

$$S(\theta) = \left[\text{diag}\ \left[e_i(\theta) \right] \right] \quad ; \quad m = n$$

$$S(\theta) = \left[\begin{array}{c} \text{diag}\ \left[e_i(\theta) \right] \\ \hline 0 \end{array} \right] \quad ; \quad m > n$$

where $e_i(\theta)$ are the invarient polynomials over R of $A(\theta)$ given by

$$e_i(\theta) = \frac{d_i(\theta)}{d_{i-1}(\theta)} \qquad (i-1,2,\ldots,\min(m,n))$$

where $d_o(\theta) \triangleq 1$ and $d_i(\theta)$ is the greatest common divisor (g.c.d.) of all the ith order minors of $A(\theta)$.

Example (3-1)

Given

$$A(\lambda) = \begin{bmatrix} 1 & \lambda-1 & \lambda+2 \\ \lambda & \lambda^2 & \lambda^2+2\lambda \\ \lambda-2 & \lambda^2-3\lambda+2 & \lambda^2+\lambda-3 \end{bmatrix}$$

then, one can compute the following

$d_0(\lambda) \triangleq 1$,

$d_1(\lambda) = 1$,

$d_2(\lambda) = 1$,

$d_3(\lambda) = \lambda(\lambda+1)$,

and the invariant factors of $A(\lambda)$ are as follows:

54

$$e_1(\lambda) = \frac{1}{1} = 1,$$

$$e_2(\lambda) = \frac{1}{1} = 1,$$

and

$$e_3(\lambda) = \frac{\lambda(\lambda+1)}{1} = \lambda(\lambda+1).$$

Thus the Smith form $S(\lambda)$ associated with $A(\lambda)$ is

$$S(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda(\lambda+1) \end{bmatrix}$$

Example (3-2)

For

$$A(s,z) = \begin{bmatrix} s & 0 & 1 \\ 0 & sz+1 & 1 \\ 0 & 0 & z \end{bmatrix} ,$$

$$d_1(s,z) = d_2(s,z) = 1 \ , \ d_3(s,z) = sz(sz+1).$$

Thus the Smith form associated with $A(s,z)$ is

$$S(s,z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & sz(sz+1) \end{bmatrix}$$

Example (3-3)

Given

$$A(s,z) = \begin{bmatrix} s+1 & 1+z(s+1) & 0 & (s+1)z & z^2 \\ s & sz+1 & -(s+1)(s+z) & sz & z^2 \\ o & s(s+1) & (s+1) & s(s+1) & sz^2(s+1) \\ s+1 & 1+z(s+1) & -(s+1)(s+2) & (s+1)z & z^2 \end{bmatrix}$$

then, one can compute the following $d_1(s,z) =$
$d_2(s,z) = 1$, $d_3(s,z) = 1(s+1)$, and $d_4(s,z) = s(s+1)^2(s+z)$.
Thus, the diagonal elements are

$$e_1(s,z) = e_2(s+z) = 1,$$

$$e_3(s,z) = \frac{(s+1)}{1} = (s+1),$$

and

$$e_4(s,z) = \frac{S(s+1)^2(s+z)}{(s+1)} = S(s+1)(s+z).$$

Finally, the Smith form $S(s,z)$ of $A(s,z)$ is

$$S(s,z) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & (s+1) & 0 & 0 \\ 0 & 0 & 0 & S(s+1)(s+z) & 0 \end{bmatrix}.$$

Equivalence of a Matrix over R to Its Smith Form

A matrix $A(\theta) \in C^{n \cdot n}(\theta)$ is said to be unimodular if $|A(\theta)|$ is nonzero complex number. It is clear that any unimodular $A(\theta) \in C^{n \cdot n}(\theta)$ has inverse $A^{-1}(\theta)$ which belongs to $C^{n \cdot n}(\theta)$. Let $A(\theta)$ and $B(\theta)$ be matrices which belong to $C^{m \cdot n}(\theta)$, then $A(\theta)$ and $B(\theta)$ are said to be equivalent over R if and only if there exists two unimodular matrices $P(\theta) \in C^{m \cdot m}(\theta)$ and $Q(\theta) \in C^{n \cdot n}(\theta)$ such that:

$$P(\theta) \, A(\theta) \, Q(\theta) = B(\theta). \tag{3-2}$$

It is easy to show that equivalent matrices over $R = C\left[\theta_1, \theta_2, \ldots, \theta_n\right]$ have the same Smith form over R. The converse of this fact is true in the case of matrices over the field R in one variable since $A(\theta) \in C^{m \cdot n}(\theta)$ in one variable is always equivalent to its Smith form.

Although it is true that two matrices $A(\theta)$, $B(\theta)$ are equivalent over $S \doteq R\left[z\right]$ (ring of polynomial in variable) if and only if they have the same Smith form, this may not be true for the case of matrices over $R = C\left[\theta_1, \theta_2, \ldots \theta_n\right]$ (polynomials of more than one variable). The following two examples illustrate this fact over the field $R = C\left[s, z\right]$ (see Frost and Storey   ).

Example (3-4)

The following matrices

$$A(s,z) = \begin{bmatrix} s+z & 0 & z \\ 0 & s+z & 0 \\ 0 & 0 & s \end{bmatrix}, \quad B(s,z) = \begin{bmatrix} s+z & 0 & 1 \\ 0 & s+z & 0 \\ 0 & 0 & s \end{bmatrix}$$

have the same Smith form $S(s,z)$

$$S(s,z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & s+z & 0 \\ 0 & 0 & s(s+z) \end{bmatrix}$$

but there is no transformation of equivalence over $R = C[s,z]$ such that (3.2) holds. Thus, $A(s,z)$ and $B(s,z)$ are not equivalent. This is clear since they have different rank at $s=z=0$.

Example (3-5):

The following matrices are not equivalent although they have the same Smith form:

$$A(s,z) = \begin{bmatrix} s & 0 & 1 \\ 0 & sz+1 & 1 \\ 0 & 0 & z \end{bmatrix}, \quad B(s,z) = \begin{bmatrix} s & 0 & 0 \\ 0 & sz+1 & 1 \\ 0 & 0 & z \end{bmatrix}$$

Their Smith form is

$$S(s,z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & SZ(sz+1) \end{bmatrix}$$

Frost and Storey $\begin{bmatrix} 8 \end{bmatrix}$, $\begin{bmatrix} 9 \end{bmatrix}$ investigated the sufficient and necessary conditions for a matrix $A(\theta) \in C^{m \cdot n}(\theta)$ in two variables to be equivalent to its Smith form. These conditions involve a new concept called zeros of a matrix over R.

Zeros of a Matrix over R

For a matrix over $R\begin{bmatrix} s \end{bmatrix}$ it is certainly the case that if the determinantal divisor $d_i(s)$ is removed from all the ith-order minors, then the remaining polynomials cannot be simultaneously zero for any value of S. This result does not extend for matrices over $R\begin{bmatrix} s,z \end{bmatrix}$.

It is quite possible that on removal of the determinantal divsor $d_i(s,z)$ from all the ith-order minors of a matrix A(s,z) over $R\begin{bmatrix} s,z \end{bmatrix}$, the remaining polynomials may all be simultaneously zero for one or more values of the pair (s,z). Such value of (s,z) will be called an ith-order zero of A(s,z).

Example (3.6):

Let us consider the same matrices given in Example (3.4). $A(s,z)$ has zeros of first and second order at $(0,0)$ while $B(s,z)$ has no zero of any order.

Example (3.7):

Let us consider the same matrices given in (3.5). $A(s,z)$ has no zero of any order while $B(s,z)$ has zero of second order at $(0,0)$.

It is important to note that the Smith form $S(s,z)$ has no zero of any order. To prove that, it is quite sufficient to note that the determinant of ith principal minor is $d_i(s,z)$ as follows:

determinant of ith principal minor

$$= e_1(s,z) \cdot e_2(s,z),\dots e_i(s,z) = \frac{d_1(s,z)}{1} \cdot \frac{d_2(s,z)}{d_1(s,z)},\dots,$$

$$\dots \cdot \frac{d_i(s,z)}{d_{i-1}(s,z)}$$

$$= d_i(s,z).$$

The Conditions for Equivalence Over $R\left[s,z\right]$

It is important to note that a transformation of equivalence over $R\left[s,z\right]$ preserves the zeros of a matrix over $R\left[s,z\right]$. From this remark it is clear that matrices over $R\left[s,z\right]$ having the same Smith form over $R\left[s,z\right]$ but not having the same zeroes are not equivalent over

60

$R[s,z]$. Applying this fact to matrices $A(s,z)$ and $B(s,z)$ given in example (3.4), it is clear that both matrices $A(s,z)$ and $B(s,z)$ do not have the same zeros. Since Smith form over $R[s,z]$ has no zeros, it is clear that a matrix over $R[s,z]$ which has zeros cannot be equivalent to its Smith form over $R[s,z]$.

As an example, note that the following matrices

$$A(s,z) = \begin{bmatrix} s+z & 0 & z \\ 0 & s+z & 0 \\ 0 & 0 & s \end{bmatrix} , \quad B(s,z) = \begin{bmatrix} s & 0 & 0 \\ 0 & sz+1 & 1 \\ 0 & 0 & z \end{bmatrix}$$

are not equivalent to their Smith form because $A(s,z)$ has zero of order 1 and 2 at (0,0); and $B(s,z)$ has a zero of second order at (0,0).

An example of matrices over $R[s,z]$ which are equivalent to their Smith form is as follows:

$$A(s,z) = \begin{bmatrix} s+z & 0 & 1 \\ 0 & s+z & 0 \\ 0 & 0 & s \end{bmatrix} , \quad B(s,z) = \begin{bmatrix} s & 0 & 1 \\ 0 & sz+1 & 1 \\ 0 & 0 & z \end{bmatrix}$$

Note that both matrices have no zeros of any order.

Lee and Žak[19] proved that lack of zeros of any order is not a sufficient condition for the equivalence of $A(s,z)$ over $R[s,z]$ to its Smith form. For example the matrix

$$A(s,z) = \begin{bmatrix} z & -s-1 \\ -s^2 & z \end{bmatrix}$$

has Smith form

$$S(s,z) = \begin{bmatrix} 1 & 0 \\ 0 & z^2-s^2(s+1) \end{bmatrix},$$

but A(s,z) is not equivalent to S(s,z) although A(s,z)
has no zeroes of any order.

Construction of Transformations of Equivalence

In the previous part we have investigated the
equivalence of a matrix to its Smith form over the ring
of polynomials in several variables. The question now
is how to construct transformation of the form (3,2) which
reduces the matrix $A(\theta) \in C^{m \cdot n}(\theta)$ to its Smith form S(θ).
That is, to find unimodular matrices P(θ) and Q(θ) such
that

$$P(\theta) \ A(\theta) \ Q(\theta) = S(\theta) \tag{3-3}$$

As we did in Chapter II, P(θ) and Q(θ) can be con-
structed using elementary row and column operations, if one
modifies the standard definitions of elementary operations
to include:

1) Multiplying of a row (column) by an arbitrary
nonzero scalar constant,

2) adding to a row (column) the elements of another
row (column) multiplied by an arbitrary polynomial which
belongs to the ring $R[\theta_1, \theta_2, \ldots, \theta_n]$, and

3) interchanging of two rows (columns).

The construction of matrices $P(\theta)$ and $Q(\theta)$ is, in
general, a difficult step. Frost and Storey
suggested a systematic procedure to reduce matrix $A(s,z)$
over the ring $R[s,z]$ to its equivalent Smith form using
elementary transformations. The first step in this pro-
cedure is to bring the matrix $A(s,z)$ via a transforma-
tion of equivalence over $R[s,z]$ to the form

$$e_1(s,z) \; \hat{A}(s,z)$$

where $e_1(s,z)$ is the first invariant polynomial of
$A(s,z)$, and $\hat{A}(s,z)$ has the form

$$\hat{A}[s,z] = \begin{bmatrix} 1 & 0 \\ & \\ 0 & A_1(s,z) \end{bmatrix} .$$

It follows that $A(s,z)$ is equivalent over $R[s,z]$ to
a matrix of the form

$$\begin{bmatrix} e_1(s,z) & 0 \\ 0 & e_1(s,z)A_1(s,z) \end{bmatrix} .$$

63

By repeating this procedure for the matrix $e_1(s,z)A_1(s,z)$, we can find a transformation which reduces $e_1(s,z)A_1(s,z)$ to the form

$$e_2(s,z) \quad A''(s,z) \ ,$$

where $A''(s,z)$ has the form

$$A''(s,z) \quad = \quad \begin{bmatrix} 1 & 0 \\ & \\ 0 & A_2(s,z) \end{bmatrix} \ .$$

That means $A(s,z)$ is equivalent over $R\begin{bmatrix} s,z \end{bmatrix}$ to

$$\begin{bmatrix} e_1(s,z) & 0 & 0 \\ 0 & e_2(s,z) & 0 \\ 0 & 0 & e_2(s,z)A_2(s,z) \end{bmatrix} \ .$$

This procedure can be successfully repeated until $A(s,z)$ has been brought to its Smith form under condition $A(s,t)$ is equivalent to its Smith form. This procedure will be explained in the next example.

Example (3-8):

Given

64

$$A(s,z) = \begin{bmatrix} s+1 & 1+z(s+1) & 0 & (a+1)z & z^2 \\ S & sz+1 & -(s+1)(s+z) & sz & z^2 \\ 0 & s(s+1) & S+1 & s(s+1) & sz^2(s+1) \\ s+1 & 1+z(s+1) & -(s+1)(s+z) & (s+1)z & z^2 \end{bmatrix}$$

one can compute

$$e_1(s,z) = 1$$

$A(s,z)$ is brought using

$$P_1 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -s & 1+s & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \quad,$$

and

$$Q_1 = \begin{bmatrix} 1 & -z & -(s+1)(s+z) & -z & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

to the form

$$\begin{bmatrix} 1 & 0 \\ 0 & A_1(s,z) \end{bmatrix}$$

where

$$A_1(s,z) = \begin{bmatrix} 1 & -(s+1)^2(s+z) & 0 & z^2 \\ s(s+1) & s+1 & s(s+1) & sz^2(s+1) \\ 0 & -(s+1)(s+z) & 0 & 0 \end{bmatrix}$$

Now $e_2(s,z) = 1$.

Again, $A_1(s,z)$ is brought using

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ -s(s+1) & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$Q_2 = \begin{bmatrix} 1 & (s+1)^2(s+z) & 0 & -z^2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

to the form

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \; A_2(s,z) \end{bmatrix}$$

where

66

$$1 \quad A_2(s,z) = \begin{bmatrix} (s+1) + (1+s(s+1)^2(s+z)) & s(s+1) & 0 \\ \\ -(s+1)(s+z) & 0 & 0 \end{bmatrix}$$

Then, using

$$P_3 = \begin{bmatrix} 1 & s(s+1)^2 \\ \\ (s+z) & 1+s(s+1)^2(s+z) \end{bmatrix} \qquad ,$$

and

$$Q_3 = \begin{bmatrix} 1 & -s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A_2 is brought to the form

$$\begin{bmatrix} s+1 & 0 & 0 \\ 0 & s(s+z)(s+1) & 0 \end{bmatrix}$$

Finally, $A(s,z)$ is reduced to the Smith form

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & s+1 & 0 & 0 \\ 0 & 0 & 0 & s(s+1)(s+z) & 0 \end{bmatrix}$$

Using

$$P(s,z) = \left[\begin{array}{c|c} I_2 & O \\ \hline O & P_3 \end{array}\right] \left[\begin{array}{c|c} 1 & O \\ \hline O & P_2 \end{array}\right] \quad P_1$$

and

$$Q(s,z) = \left[\begin{array}{c|c} I_1 & O \\ \hline O & Q_3 \end{array}\right] \left[\begin{array}{c|c} 1 & O \\ \hline O & Q_2 \end{array}\right] \quad Q_1$$

Generalized Inverses of Polynomial Matrices of Constant Rank

The main purpose of this section is to establish an algorithm to compute different generalized inverses of a given matrix $A(\theta) \in C^{m \cdot n}(\theta)$. Throughout this section, it is assumed that a matrix $A(\theta)$ is equivalent to its Smith form; i.e., to the following form:

$$\left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array}\right] , \tag{3-4}$$

where $\Omega(\theta)$ is a diagonal matrix which belongs to $C^{(m-r)(n-r)}(\theta)$. That is, there exist two unimodular matrices $P(\theta)$, $Q(\theta)$ such that

$$P(\theta) \ A(\theta) \ Q(\theta) = \begin{bmatrix} I_r & O \\ \hline O & \Omega(\theta) \end{bmatrix} \qquad (3-5)$$

The next theorem establishes the necessary and sufficient conditions for a matrix $A(\theta) \in C^{m \cdot n}(\theta)$ to have generalized inverses $A_{1,2}(\theta)$, $A_{1,2,3}(\theta)$,...., etc.

<u>Theorem (3.1)</u>: Sontag $\begin{bmatrix} 27 \end{bmatrix}$

Let $A(\theta) \in C^{m \cdot n}(\theta)$, then $A(\theta)$ has $A_{1,2}$ - inverse which belongs to $C^{n \cdot m}(\theta)$ if and only if $A(\theta)$ has a constant rank $r \leq \min(m,n)$.

.<u>Proof</u>: Suppose that $A(\theta)$ has constant rank r, then $A(\theta)$ can be reduced to the form (3-1) using elementary transformation as follows:

$$P(\theta) \ A(\theta) \ Q(\theta) = \begin{bmatrix} I_r & O \\ \hline O & O \end{bmatrix} \qquad (3-6)$$

where $P(\theta)$ and $Q(\theta)$ are unimodular matrices, let

$$X = Q(\theta) \begin{bmatrix} I_r & O \\ \hline O & O \end{bmatrix} \ P(\theta)$$

It is clear that X is an $A_{1,2}(\theta)$ inverse of $A(\theta)$.

Remark

By proper partitioning of the matrices $P(\theta)$ and $Q(\theta)$ given in (3-6) in the same way as in the case of constant matrices, i.e.,

$$P(\theta) = \left[\frac{T(\theta)}{M(\theta)}\right] \qquad ; \qquad Q(\theta) = \left[S(\theta) \mid N(\theta)\right]$$

then,

$$A_{1,2}(\theta) = S(\theta)\ T(\theta)$$

Remark

In the case of full raw rank matrix $A(\theta) \epsilon\ C^{m \cdot n}(\theta)$, i.e., rank of $A(\theta) = m$

$$A_{1,2}(\theta) = A_{1,2,3}(\theta)$$

Remark

In the case of full column rank matrix $A(\theta) \epsilon\ C^{m \cdot n}(\theta)$, i.e., rank of $A(\theta) = \dot{n}$

$$A_{1,2}(\theta) = A_{1,2,4}(\theta)$$

Theorem (3-2):

Let $A(\theta)$ be a polynomial matrix which belongs to $A^{m \times n}(\theta)$. Let $P(\theta)$ and $Q(\theta)$ be unimodular matrices such that

$$P(\theta) \ A(\theta) \ Q(\theta) = \left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array} \right]$$

where $\Omega(\theta) \in C^{(m-r) \cdot (n-r)}(\theta)$, then there exists $A_2(\theta)$ with rank $k \leq r$ in the following form

$$A_2(\theta) = Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array} \right] P(\theta). \tag{3-7}$$

By proper partitioning of $P(\theta)$, $Q(\theta)$ as follows

$$P(\theta) = \left[\frac{T(\theta)}{M(\theta)} \right] \quad , \quad A(\theta) = \left[S(\theta) \mid N(\theta) \right] \tag{3-8}$$

where $T(\theta) \in C^{k \cdot m}(\theta)$, $M(\theta) \in C^{(m-k) \cdot m}(\theta)$, $S(\theta) \in C^{n \cdot k}(\theta)$, and $N(\theta) \in C^{n \cdot (n-k)}$,

$A_2(\theta)$ can be written as

$$A_2(\theta) = S(\theta) \ T(\theta) \tag{3-9}$$

<u>Proof</u>:

$$A_2(o) \ A(\theta) \ A_2(\theta) = Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] P(\theta) \ P^{-1}(\theta) \left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array}\right]^{Q-1}(\theta)$$

$$Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] \ P(\theta)$$

$$= Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] \left[\begin{array}{c|c} I_k & O \\ \hline O & \begin{array}{c} I_{r-k} \\ \Omega(\theta) \end{array} \end{array}\right] \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] \ P(\theta)$$

$$= Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] P(\theta) = A_2(\theta)$$

This proves the first part; to prove the second part, substitute (3-8) into the relation (3-7) as follows

$$A_2(\theta) = Q(\theta) \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] \ P(\theta)$$

$$= \left[S(\theta) \ \middle| \ N(\theta)\right] \left[\begin{array}{c|c} I_k & O \\ \hline O & O \end{array}\right] \left[\begin{array}{c} T(\theta) \\ M(\theta) \end{array}\right] = \left[S(\theta) \ | \ O\right] \left[\begin{array}{c} T(\theta) \\ M(\theta) \end{array}\right]$$

$$= S(\theta) \ T(\theta) \ .$$

In the following part we will demonstrate the existence of different types of generalized inverses over $R = C \left[\theta_1, \theta_2, \ldots, \theta_n\right]$.

<u>Example (3 -9)</u>:

Given

72

$$A(s,z) = \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \quad ,$$

then one can construct the following array using elementary transformation

$$\left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & s^2 \\ 0 & 1 & 0 & 0 & 0 & 1 & -sz \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ \hline 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & 0 & \\ 1 & 0 & 0 & -sz & & & \\ 0 & 1 & 0 & z^2-1 & & & \end{array}\right] \triangleq \left[\begin{array}{c|c|c} I_r & 0 & T(\theta) \\ \hline S(\theta) & N(\theta) & 0 \end{array}\right]$$

The following matrix is an $A_{1,2,3}$-inverse of $A(\theta)$.

$$A_{1,2,3}(\theta) = S(\theta)\,T(\theta) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & s^2 \\ 0 & 1 & -sz \\ 0 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix}$$

One can check the answer by substituting in the first
three axioms (   ) as follows:

$$A \; A_{1,2,3} A = \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix} \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & sz & 1 & 0 \\ 0 & 1-z^2 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \quad = A$$

also

$$A_{1,2,3} \; A \; A_{1,2,3} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix} \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & sz & 1 & 0 \\ 0 & 1-z^2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & sz \\ 0 & 1 & -sz \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix} = A_{1,2,3}$$

74

and finally

$$A \, A_{1,2,3} = \begin{bmatrix} s^2 & sz & 1 & 0 \\ -sz & 1-z^2 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & -sz \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad .$$

Also, we compute $A_2(\theta)$ with rank = 1 as follows:

$$A_2(\theta) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & s^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 0 & 0 \end{bmatrix}$$

and finally we compute $A_2(\theta)$ with rank = 2 as follows:

$$A_2(\theta) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & s^2 \\ 0 & 1 & -sz \end{bmatrix}$$

To compute $A_{1,2,34}(\theta)$, it is necessary to satisfy the relation

$$\tilde{S}^t \, N = 0.$$

Example (3-10)

Given

$$A = \begin{bmatrix} s+z & 0 & 1 \\ 0 & s+z & 0 \\ 0 & 0 & s \end{bmatrix}$$

then, one can construct the following array

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & s+z & 0 & 0 & 1 & 0 \\ 0 & 0 & s(s+z) & S & 0 & -1 \\ \hline 0 & 0 & 1 & & & \\ 0 & 1 & 0 & & & \\ 1 & 0 & -(s+z) & & & \end{array}\right]$$

It is the case of nonconstant rank so we can construct only $A_2(s,z)$ with rank = 1 as follows

$$A_2(s,z) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Example (3-11)

Given

$$A(s,z) = \begin{bmatrix} S & 1-z^2 & 0 & z \\ 0 & S & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

then one can construct the following array

$$\left[\begin{array}{ccc|c||ccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & -s \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & -z & & & \\ 0 & -s & 1 & sz & & & \\ 0 & z & 0 & 1-z^2 & & & \end{array}\right]$$

Thus $A_{1,2,3}$ is

$$A_{1,2,3}(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -s & 1 \\ 0 & z & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -s \\ - & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -s \\ -s & 1 & s^2 \\ z & 0 & -sz \end{bmatrix}$$

and $A_2(\theta)$ with rank $= 2$ is

$$A_2(\ ) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -s \\ 0 & z \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -s \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -s \\ -s & 0 & s^2 \\ z & 0 & -sz \end{bmatrix}$$

Example (3-12)

For the matrix

$$A(s,z) = \begin{bmatrix} S & Z & S^2+Z+1 \\ S^2Z & SZ^2 & S^3Z+SZ^2+SZ \end{bmatrix}$$

we can construct the array

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & -SZ & 1 \\ \hline -S & SZ & 1+S^2 & & \\ -1 & 1+Z & Z & & \\ 1 & -Z & -S & & \end{array} \right]$$

$$A_{1,2}(Z) = \begin{bmatrix} -S \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -S & 0 \\ -1 & 0 \\ 1 & 0 \end{bmatrix}$$

The Matrices with Variable Rank

For $A(\theta) \in C^{m \cdot n}(\theta)$ and $A(\theta)$ does not have constant rank, the problem of finding solutions of $A(\theta) X(\theta) = b(\theta)$

arises since it was not treated by Sontag $\begin{bmatrix} 27 \end{bmatrix}$ and others. Throughout the next section it will be assumed that there exist unimodular matrices $P(\theta)$, $Q(\theta)$ belonging to $C^{m \cdot m}(\theta)$, $C^{n \cdot n}(\theta)$, respectively, such that:

$$P(\theta) \, A(\theta) \, Q(\theta) = \left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array} \right] = A_o(\theta) \qquad (3\text{-}10)$$

where $A_o(\theta)$ is the Smith form of $A(\theta)$. Such matrices were treated by Frost and Storey $\begin{bmatrix} 8,9 \end{bmatrix}$, and Lee and Zak $\begin{bmatrix} 19 \end{bmatrix}$, when matrices were reduced to their equivalent Smith form:

Theorem(3-3):

Let $A(\theta) \; \varepsilon \; C^{m \cdot n}(\theta)$, $b(\theta) \; \varepsilon \; C^{m \cdot 1}(\theta)$, and $A(\theta)$ has nonconstant rank $k(\theta)$, i.e.,

$$1 \; \leq \; r \leq \; k(\theta) \; \leq \min(m,n).$$

Let $P(\theta)$, $Q(\theta)$ be unimodular matrices such that:

$$P(\theta) \, A(\theta) \, Q(\theta) = \left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array} \right] \quad ,$$

then the following partitioning of matrices

$$\left[\begin{array}{c|c} A(\theta) & I_m \\ \hline I_n & O \end{array} \right] \quad , \quad \left[\begin{array}{c|c|c} I_r & O & T(\theta) \\ \hline O & \Omega(\theta) & M(\theta) \\ \hline S(\theta) & N(\theta) & O \end{array} \right] \qquad (3\text{-}11)$$

are equivalent over $R = C \begin{bmatrix} \theta_1, \ldots, \theta_n \end{bmatrix}$

Proof: Proof is similar to constant case of A.

Theorem (3-4):

Let $A(\theta) \in C^{m \cdot n}(\theta)$, and $A(\theta)$ have nonconstant rank, i.e.,

$$1 \leq r \leq \text{rank of } A(\theta) \leq \min(n,m),$$

and the hypothesis of theorem (3-3) holds.

Then the following set of equations

$$A(\theta) X(\theta) = b(\theta) \tag{3-12}$$

has a solution $X(\theta) \in C^{n \cdot 1}(\theta)$ if

$$M(\theta) b(\theta) = \Omega(\theta) Z(\theta), \tag{3-13}$$

for some $Z(\theta) \in C^{(n-r) \cdot 1}(\theta)$, and in which case the general solution $X(\theta)$ is given by

$$X(\theta) = S(\theta) T(\theta) b(\theta) + N(\theta) Z(\theta) \tag{3-14}$$

where $S(\theta)$, $T(\theta)$, $N(\theta)$, $\Omega(\theta)$ are given as in (3-11)

Proof: For any $A(\theta) \in C^{m \cdot n}(\theta)$ there exist unimodular matrices such that

$$P(\theta) A(\theta) Q(\theta) = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & \Omega(\theta) \end{array}\right] \tag{3-15}$$

where $P(\theta) \in C^{m \cdot m}(\theta)$, and $Q(\theta) \in C^{n \cdot n}(\theta)$. $A(\theta) X(\theta) = b(\theta)$

80

has a solution $X(\theta)$ iff

$P(\theta)\, A(\theta)\, X(\theta) = P(\theta)\, b(\theta)$ has a solution $X(\theta)$ iff

$(P(\theta)\, A(\theta)\, Q(\theta))\, Q^{-1}(\theta)\, X(\theta) = P(\theta)\, b(\theta)$ has a solution $X(\theta)$ iff

$$\left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array}\right] y(\theta) = P(\theta)\, b(\theta)$$ has a solution $y(\theta) = Q^{-1}(\theta)\cdot X(\theta)$ iff

$$\left[\begin{array}{c|c} I_r & O \\ \hline O & \Omega(\theta) \end{array}\right]\left[\frac{W(\theta)}{Z(\theta)}\right] = \left[\frac{T(\theta)}{M(\theta)}\right] b(\theta) \; ; \quad P(\theta) = \left[\frac{T(\theta)}{M(\theta)}\right] \; ;$$

$$y(\theta) = \left[\frac{W(\theta)}{Z(\theta)}\right] \; ; \quad X(\theta) = Q(\theta)\, y(\theta) = \left[S(\theta)\,\middle|\,N(\theta)\right]\left[\frac{W(\theta)}{Z(\theta)}\right] \; .$$

The last set of equations can be written as

$W(\theta) = T(\theta)\, b(\theta) \; ; \quad \Omega(\theta)\, Z(\theta) = M(\theta)\, b(\theta) \; ;$

$X(\theta) = S(\theta)\, W(\theta) + N(\theta)\, Z(\theta).$

Thus, the solution of $A(\theta)\, X(\theta) = b(\theta)$ will be:

$X(\theta) = S(\theta)\, T(\theta)\, b(\theta) + N(\theta)\, Z(\theta)$

in the condition that

$\Omega(\theta)\, Z(\theta) = M(\theta)\, b(\theta)$, holds for some $Z(\theta) \in C^{(n-r)\cdot 1}(\theta)$

with appropriate size.

Example (3-13)

Consider the following system

$$\begin{bmatrix} 1-\lambda & S & \lambda & 0 \\ Z & SZ-\lambda & 0 & \lambda \end{bmatrix} X(S,\lambda,Z) = \begin{bmatrix} \lambda(S+Z) \\ \lambda^2 + SZ\lambda \end{bmatrix}$$

Using the elementary operations we can construct
the following array:

$$
\left[
\begin{array}{c|ccc||cc}
1 & 0 & 0 & 0 & 1 & 0 \\
\hline
0 & \lambda & 0 & 0 & -Z & 1 \\
\hline\hline
1 & 0 & -\lambda & -S & & \\
0 & 0 & 0 & 1 & & 0 \\
1 & 0 & 1-\lambda & -S & & \\
0 & 1 & Z & 1 & &
\end{array}
\right]
$$

To check the consistency condition, one may calculate

$$
\begin{bmatrix} \lambda & 0 & 0 \end{bmatrix}
\begin{bmatrix} a(\theta) \\ d(\theta) \\ c(\theta) \end{bmatrix}
=
\begin{bmatrix} -Z & 1 \end{bmatrix}
\begin{bmatrix} \lambda(s+z) \\ \lambda^2 + SZ\lambda \end{bmatrix}
$$

that is

$$\lambda\, a(\theta) = -Z\lambda(s+z) + \lambda^2 + SZ\lambda, \text{ and } d(\theta)\ ,\ c(\theta) \text{ are}$$

arbitrary polynomial.

$$
Z(\theta) =
\begin{bmatrix} -Z(s+z)+\lambda+SZ \\ d(\theta) \\ c(\theta) \end{bmatrix}
=
\begin{bmatrix} \lambda - z^2 \\ d(\theta) \\ c(\theta) \end{bmatrix}
$$

The general solution will be

$$X = S \, T \, b + NZ$$

$$= \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda(s+z) \\ \lambda^2 + SZ\lambda \end{bmatrix} + \begin{bmatrix} 0 & -\lambda & -S \\ 0 & 0 & 1 \\ 0 & 1-\lambda-S \\ 1 & Z & 1 \end{bmatrix} \begin{bmatrix} \lambda - z^2 \\ d(\theta) \\ c(\theta) \end{bmatrix}$$

$$= \lambda(s+z) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\lambda d(\theta) \; -S \, c(\theta) \\ c(\theta) \\ (1-\lambda)d(\theta) \; -S \, c(\theta) \\ \lambda \cdots z^2 + z \; d(\theta) + c(\theta) \end{bmatrix}$$

$$= \lambda(s+z) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + d(\theta) \begin{bmatrix} -\lambda \\ 0 \\ 1-\lambda \\ z \end{bmatrix} + c(\theta) \begin{bmatrix} -S \\ 1 \\ -S \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \lambda - z^2 \end{bmatrix}$$

choosing $d(\theta) = Z$ $c(\theta) = -\lambda$

$$x = \lambda(s+z) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\lambda z \\ 0 \\ z - \lambda z \\ z^2 \end{bmatrix} + \begin{bmatrix} \lambda S \\ -\lambda \\ \lambda S \\ -\lambda \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \lambda - z^2 \end{bmatrix}$$

$$= \begin{bmatrix} 2\lambda S \\ -\lambda \\ 2\lambda s + z \\ 0 \end{bmatrix}$$

83

$$Ax = \begin{bmatrix} 1-\lambda & S & \lambda & 0 \\ Z & SZ-\lambda & 0 & \lambda \end{bmatrix} \begin{bmatrix} 2\lambda S \\ -\lambda \\ 2\lambda s+z \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda S+\lambda Z \\ \lambda^2+\lambda SZ \end{bmatrix} = b$$

IV Riccati and Lyapunov Matrix Equation

The main purpose of this part is to establish
methods to solve algebraic Riccati and Lyapunov equa-
tions. The Lyapunov equation

$$AX-XB = C \qquad (4-1)$$

and the Riccati matrix equation

$$AX-XB + xDx = C \qquad (4-2)$$

where the matrices A, B, C, and D have elements which belong
to the field of complex numbers.

Riccati and Lyapunov equations are very important
to establish methods to enable any one to do systems
decomposition, i.e., transform large systems to uncoupled
small subsystems. Such process requires the solution
of the equations (4-1) or (4-2).

At first the elements of all matrices are real or
complex numbers. The notion of strong similarity of the
following pair of matrices:

$$\left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right] \qquad (4-3)$$

and

85

$$\left[\begin{array}{c|c} A & O \\ \hline O & B \end{array}\right] \tag{4-4}$$

whenever the equation (4-1) holds and

$$\left[\begin{array}{c|c} I & X \\ \hline O & I \end{array}\right] \left[\begin{array}{c|c} A & C \\ \hline O & B \end{array}\right] \left[\begin{array}{c|c} I & -X \\ \hline O & I \end{array}\right] = \left[\begin{array}{c|c} A & -AX+XB+C \\ \hline O & B \end{array}\right] = \left[\begin{array}{c|c} A & O \\ \hline O & B \end{array}\right] \tag{4-5}$$

Equation (4-5) holds whenever the equation $-AX+XB+C=0$

has a solution (i.e., equation (4-1)).

For applying to systems decomposition consider the
following differential system of equations.

$$\frac{dx(t)}{dt} = \left[\begin{array}{c|c} A & C \\ \hline O & B \end{array}\right] x(t) \tag{4-6}$$

let

$$x(t) = \left[\begin{array}{c|c} I & -X \\ \hline 0 & I \end{array}\right] y(t)$$

where

$$AX - XB = C \tag{4-7}$$

This system is reduced to

$$\frac{dy(t)}{dt} = \left[\begin{array}{c|c} I & X \\ \hline C & I \end{array}\right] \left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right] \left[\begin{array}{c|c} I & -X \\ \hline 0 & I \end{array}\right] y(t) = \left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right] y(t) \tag{4-8}$$

So the systems (4-6) and (4-3) are strongly similar under the condition that (4-1) holds. System (4-8) is uncoupled sub-systems.

Now consider the fully coupled differential system of equations

$$\frac{dx(t)}{dt} = \left[\begin{array}{c|c} B & D \\ \hline C & A \end{array}\right] x(t) \tag{4-9}$$

applying the following transformation to (4-9)

$$x(t) = \left[\begin{array}{c|c} I & 0 \\ \hline -X & I \end{array}\right] y(t) \tag{4-10}$$

where X is a solution of (4-2) with matrices coefficient defined as in (4-9).

$$\frac{dy(t)}{dt} = \left[\begin{array}{c|c} I & 0 \\ \hline X & I \end{array}\right] \left[\begin{array}{c|c} B & D \\ \hline C & A \end{array}\right] \left[\begin{array}{c|c} I & 0 \\ \hline -X & I \end{array}\right] y(t) \tag{4-11}$$

$$= \left[\begin{array}{c|c} B & D \\ \hline XB+C & XD+A \end{array}\right] \left[\begin{array}{c|c} I & C \\ \hline X & I \end{array}\right] y(t)$$

$$= \left[\begin{array}{c|c} B-DX & D \\ \hline XB+C-XDX-AX & XD+A \end{array}\right] y(t)$$

$$= \left[\begin{array}{c|c} B-DX & D \\ \hline 0 & XD+A \end{array}\right] y(t)$$

Equation (4-11) represents a partially coupled system instead of (4-9).

87

Applying another transformation as before using Lyapounov equation (4-1), the system will be reduced to

$$\frac{dz(t)}{dt} = \begin{bmatrix} B-DX & 0 \\ 0 & A+XD \end{bmatrix} z(t)$$

In the next two parts we will establish the different techniques and approaches used to solve equations (4-1) and (4-2).

Solving Lyapunov Equation

Consider the linear system represented by

$$\frac{dx}{dt} = Rx \qquad (4-12)$$

where

$$R = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \qquad (4-13)$$

where A, B, C are n·n matrices whose elements are complex numbers. The matrix R is similar to $R^* = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ whenever the equation $AX - XB = C$ has a solution X. In this case it is clear that

$$TRT^{-1} = R^* \qquad (4-14)$$

where

$$T = \begin{bmatrix} I & X \\ \hline 0 & I \end{bmatrix} \qquad (4-15)$$

where X is a solution for Lyapounov equation (4-1).

First Technique. Theorem (4-1): Roth (   )

A necessary and sufficient condition that equation (4-1) has a solution X where the matrices A, B, C are square matrices of order $n \cdot n$ with the elements in the field of complex numbers is that the matrices R and R* given in (4-13) are similar.

Letting $f_A(\lambda)$ and $f_B(\lambda)$ be the characteristic polynomial of A,B respectively then

$$f_A(R) = \begin{bmatrix} 0 & M \\ \hline 0 & U \end{bmatrix} , \quad f_B(R) = \begin{bmatrix} N & M \\ \hline 0 & 0 \end{bmatrix} \qquad (4-16)$$
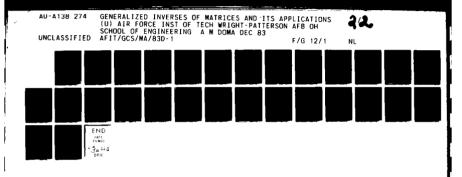
Proof:

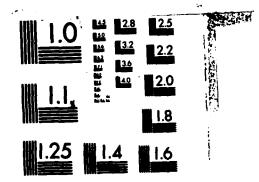The first part of the theorem is clear using the following equation

$$\begin{bmatrix} I & X \\ \hline 0 & I \end{bmatrix} \quad R \quad \begin{bmatrix} I & -X \\ \hline 0 & I \end{bmatrix} = \begin{bmatrix} A & -AX+XB+C \\ \hline 0 & B \end{bmatrix} \qquad (4-17)$$

if X is a solution of (4-1) then

$$\begin{bmatrix} I & X \\ \hline 0 & I \end{bmatrix} \cdot R \quad \begin{bmatrix} I & -X \\ \hline 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ \hline 0 & B \end{bmatrix} = R* \qquad (4-18)$$

$$R = \begin{bmatrix} I & -X \\ \hline 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ \hline 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ \hline 0 & I \end{bmatrix} \qquad (4-19)$$

89

END
DATE
FILMED
3-84
DTIC

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

$$f_A(R) = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} f_A(A) & 0 \\ 0 & f_A(B) \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \qquad (4\text{-}20)$$

$$f_A(R) = \begin{bmatrix} 0 & -Xf_A(B) \\ 0 & f_A(B) \end{bmatrix} = \begin{bmatrix} 0 & M \\ 0 & U \end{bmatrix} \qquad (4\text{-}21)$$

This implies that

$$M = -XU \qquad (4\text{-}22)$$

In the same way

$$f_B(R) = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} f_B(A) & 0 \\ 0 & f_B(B) \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} f_B(A) & f_B(A) \cdot X \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{bmatrix}$$

This implies

$$\hat{M} = \hat{N} \cdot X \qquad (4\text{-}23)$$

Theorem (4-2)

The equation (4-1) has a solution X if the following pair of equations has a common solution:

$$M + XU = 0 \qquad (4\text{-}24)$$

$$\hat{M} - \hat{N}X = 0 \qquad (4\text{-}25)$$

where M, U, $\hat{M}$ and $\hat{N}$ are as given in (4-16). Moreover, any common solution will be a solution of (4-1).

The necessary and sufficient condition that (4-1) has a solution X is that the equations

$$M\,U^-U = M \qquad (4\text{-}26)$$

and

$$\hat{N}\,\hat{N}^- \cdot \hat{M} = \hat{M} \qquad (4\text{-}27)$$

and

$$\hat{M}\,U = -\hat{N}\,M \qquad \text{hold.} \qquad (4\text{-}28)$$

In this case the solution will be expressed as

$$X = \hat{N}^- \cdot \hat{M} - M\,U^- + \hat{N}^- \cdot \hat{N}\,M\,U^- \qquad (4\text{-}29)$$

Proof:

The first part of the theorem is clear using (4-22) and (4-23) in theorem (4-1).

Equations (4-26) and (4-27) are the consistency condition of each of the pair of the equations and equation (4-28) is the condition that the two equations have a common solution.

Example (4-1): Solve the Lyapounov equation (4-1) for

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \;, \quad B = \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix} \;, \quad C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}$$

Solution:

$$R = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 3 \\ 1 & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \end{array} \right]$$

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

$$f_B(\lambda) = \lambda^2 + \lambda$$

$$f_A(R) = \left[\begin{array}{cc|cc} 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & -2 \\ \hline 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \end{array} \right] = \left[\begin{array}{c|c} 0 & M \\ \hline 0 & U \end{array} \right]$$

$$f_B(R) = \left[\begin{array}{cc|cc} 2 & 0 & 2 & 2 \\ 2 & 0 & 2 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] = \left[\begin{array}{c|c} \hat{N} & \hat{M} \\ \hline 0 & 0 \end{array} \right]$$

First, calculate generalized inverses

$$U^- = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$

$$\hat{N}^- = \begin{bmatrix} \tfrac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}$$

Check of consistency conditions

$$M \ U^- U = M$$

$$\hat{N} \ \hat{N}^- \cdot \hat{M} = \hat{M}$$

and

$$\hat{M} \ U = -\hat{N} \ M = \begin{bmatrix} 0 & -8 \\ 0 & -8 \end{bmatrix}$$

A solution is

$$X = \hat{N}^- \hat{M} - M \ U^- + \hat{N}^- \cdot \hat{N} \ M \ U^-$$

$$= \begin{bmatrix} \tfrac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 0 & -4 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \tfrac{1}{2} & 0 \end{bmatrix}$$

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

Moreover, we can find the general solution by finding the general solution for each equation. The general solution for the first equation is:

$$X_1 = 0 - M U^- + Y_1(I - U U^-) \tag{4-30}$$

$$= -\begin{bmatrix} 0 & -4 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \tfrac{1}{2} & 0 \end{bmatrix} + Y_2\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \tfrac{1}{2} & 0 \end{bmatrix}\right)$$

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

where ℓ, m are arbitrary.

$$X_2 = \hat{N}^- \cdot \hat{M} - (I - \hat{N}^- \hat{N})Y_2 \tag{4-31}$$

$$= \begin{bmatrix} \tfrac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} - \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \tfrac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}\right)Y_2$$

$$= \begin{bmatrix} 1 & 1 \\ \ell' & m' \end{bmatrix}$$

where ℓ', m' are arbitrary.

Equating $X_1 = X_2$

$$\ell = 1 \; , \; m' = m \; , \; \ell' = 1-m$$

So the general solution will be

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

The Second Technique.

Theorem (4-3): If $f_\alpha(\lambda)$, $f_\beta(\lambda)$ are polynomials of degree n of λ with coefficient in the field of complex numbers such that

$$f_\alpha(R) = \left[\begin{array}{c|c} V & N \\ \hline 0 & 0 \end{array}\right] \qquad (4\text{-}32)$$

$$f_\beta(R) = \left[\begin{array}{c|c} 0 & \hat{N} \\ \hline 0 & \hat{M} \end{array}\right] \qquad (4\text{-}33)$$

where $R = \left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right]$, if V^{-1} exists, then a solution X of $N - VX = 0$ is a solution of (4-1). Moreover, if $\hat{M}^{-1}$ exists, then a solution X of $\hat{N} + X\hat{M} = 0$ is also a solution of (4-1).

Proof: The matrices $f_\alpha(R)$ and R commute which implies

$$\left[\begin{array}{c|c} V & N \\ \hline 0 & 0 \end{array}\right]\left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right] = \left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right]\left[\begin{array}{c|c} V & N \\ \hline 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{c|c} VA & VC+NB \\ \hline 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} AV & AN \\ \hline 0 & 0 \end{array}\right]$$

This implies the following identities

$$AV = VA \ , \ AN = VC + NB \qquad (4\text{-}34)$$

95

if X is a solution of N−VX=0 , then, using (4-34), the
following holds

$$0 = A(N - VX)$$
$$= VC + NB - AVX$$
$$= V(C + XB - AX)$$

and since V^{-1} exists, X is a solution of (4-1). In
the same way, we can prove the second part.

Example (4-2)

Solve the same problem in example (4-1)

$$f(\lambda) = |R - \lambda I| = (\lambda^2 - \lambda)(\lambda^2 + \lambda)$$

All the possible polynomials are:

Case-1 $f(\lambda) = \lambda^2 - \lambda$

Case-2 $f(\lambda) = \lambda^2 + \lambda$

Case-3 $f(\lambda) = \lambda^2$

Case-4 $f(\lambda) = \lambda^2 - 1$

Case-1:

$$f(R) = \begin{bmatrix} 0 & \begin{array}{cc} 0 & -4 \\ 0 & -2 \end{array} \\ \hline 0 & \begin{array}{cc} 0 & 2 \\ 0 & 2 \end{array} \end{bmatrix} = \begin{bmatrix} 0 & \hat{N} \\ \hline 0 & \hat{M} \end{bmatrix}$$

Consider the equation

$$N + XM = 0 \quad , \quad M^{-1} \text{ does not exist, has the general}$$

solution

$$X = -N M^- + Y \ (I - M M^-) \tag{4-35}$$

$$= \begin{bmatrix} \ell^{2+\ell} \\ m \ 1 - m \end{bmatrix} \qquad \ell, \text{ m arbitrary}$$

Substituting into the equation (4-1) by this solution,
we obtain the condition on ',m' to make (4-35)a solution.

$$\ell = 1 \quad , \quad \text{m arbitrary}$$

i.e., the general solution is

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

Case-2

$$f(r) = R^2 + R$$

$$= \left[\begin{array}{cc|cc} 2 & 0 & 2 & 2 \\ 2 & 0 & 2 & 2 \\ \hline & 0 & & 0 \end{array}\right] = \left[\begin{array}{c|c} V & N \\ \hline 0 & 0 \end{array}\right]$$

Consider the equation

N = VX   , V is singular, has the general solution

X = V⁻ N + (I − V V⁻)Y

$$= \begin{bmatrix} 1 & 1 \\ \ell' & m' \end{bmatrix} \quad , \quad \ell', m' \quad \text{arbitrary}$$

Again, substituting in equation (4-1) we have the same solution.

<u>Case-3</u>

$$f(\lambda) = R^2$$

$$= \left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

The theorem cannot be applied.

<u>Case-4</u>

$$f(\lambda) = R^2 - I$$

$$= \left[\begin{array}{cc|cc} 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ \hline 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The theorem cannot be applied.

So the previous theorem is applicable only for the cases where

$$f(\lambda) \text{ is either } f_A(\lambda) \text{ or } f_B(\lambda).$$

The Third Technique

In this part, solutions for Lyapounov equation will be obtained in terms of the principal idempotent and nilpotent matrices associated with the matrices A, B.

To solve equation (4-1), the following more general equation will be considered

$$AXE + DXB = C \tag{4-36}$$

If A is an n·n matrix having elements which belong to the field of complex numbers and $\{a_i\}$ is the set distinct characteristic roots of A, then A has the following representation:

$$A = \sum_{j=1}^{m} (a_j A_j + \overline{A}_j) \qquad m \leq n \tag{4-37}$$

where the matrices $\{A_j\}$, $\{\overline{A}_i\}$ form a complete set of principal idempotent and nilpotent matrices with the following property:

$$A_i A_j = S_j A_j , \sum_{j=1}^{m} A_j = I ; j = 1,2,\ldots,m \tag{4-38}$$

99

$$A_j \bar{A}_k = 0 \quad \text{for } j \neq k , \quad A_j \bar{A}_j = \bar{A}_j A_j = \bar{A}_j \tag{4-39}$$

Theorem (4-4)

Consider the following matrix equation

$$AXE + DXB = C \tag{4-40}$$

where the matrices A, E, D, and B have the following representation in terms of a complex set of idempotent and nilpotent.

$$A = \sum_{j=1}^{ja} (a_j A_j + \bar{A}_j) ;$$

$$B = \sum_{j=1}^{jb} (b_j B_j + \bar{B}_j) ;$$

$$D = \sum_{j=1}^{jd} (d_j D_j + \bar{D}_j) ;$$

$$E = \sum_{j=1}^{je} (e_j E_j + \bar{E}_j)$$

with

$$\bar{A}_j C = C\bar{B}_k = 0$$

$$\bar{A}\, C = C\bar{E}_k = 0$$

$$\bar{D}_j C = C\bar{B}_k = 0$$

$$\bar{D}_j C = C\bar{E}_k = 0$$

for all possible j,k and whenever $a_i e_j + d_k b_e = 0$, then

$$A_i \, D_k \, CE_j \, B_e = 0$$

The equation (4-40) has a solution X given by

$$X = \sum_{e=1}^{je} \; \sum_{k=1}^{jd} \; \sum_{j=1}^{je} \; \sum_{i=1}^{ja} \; \frac{A_i D_k C \, E_j B_e}{a_i e_j + d_k b_e} \qquad (4\text{-}41)$$

Lemma (4-1)

The solution for the Lyapounov equation (4-1) given that A,B are given as in (4-41) is

$$X = \sum_{j=1}^{i_b} \; \sum_{i=1}^{i_a} \; \frac{A_i C \, B_j}{a_i - b_j}$$

Example (4-3): Solve the same problem given in Example (4-1).

101

$$A = a_1 A_1 + a_2 A_2$$

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

$$B = b_1 B_1 + b_2 B_2$$

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

Notice that $a_1 - b_1 = 0$ and

$$A_1 CB_1 = \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

the solution is given by

$$X = \frac{A_1 CB_2}{0-(-1)} + \frac{A_2 CB_1}{1-(0)} + \frac{A_2 CB_2}{1-(-1)}$$

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

$$= \tfrac{1}{2} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Example (4-4) Consider equation (4-1) for

102

$$A = \begin{bmatrix} 3 & -7 & -20 \\ 0 & -5 & -14 \\ 0 & 3 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix}$$

$$A = 1 A_1 + 2 A_2 + 3 A_3$$

$$= 1 \begin{bmatrix} 0 & -11/2 & -11 \\ 0 & 7 & 14 \\ 0 & -3 & -6 \end{bmatrix} + 2 \begin{bmatrix} 0 & 18 & 42 \\ 0 & -6 & -14 \\ 0 & 3 & 7 \end{bmatrix} + 3 \begin{bmatrix} 1 & -25/2 & -31 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

for

$$B = 2 \begin{bmatrix} 3 & 3 & 0 \\ -2 & -2 & 0 \\ -2 & -3 & 1 \end{bmatrix} + \begin{bmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{bmatrix} + 1 \begin{bmatrix} -2 & -3 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix}$$

$$= b_1 B_1 + \bar{B}_1 + b_2 B_2$$

$$- \quad \bar{B}_j = \begin{bmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{bmatrix}$$

check that

$$C \left(\sum \bar{B}_j \right) = 0$$

$a_1 = 1, a_2 = 2, a_3 = 3$, $b_1 = 2, b_2 = 1$; therefore
$a_1 - b_2 = a_2 - b_1 = 0$

and

$$A_1C\ B_2 = 0$$

$$A_2C\ B_1 = 0$$

and the solution is

$$X = A_1C\ B_1 + A_2CB_2 + \tfrac{1}{2}\ A_3CB_2 + A_3CB_1$$

$$= \begin{bmatrix} 6 & 17/2 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix}$$

Check that $X(\sum \bar{B}_j) = X(\sum \bar{B}_j) = (\sum \bar{B}_j)X = 0$

The Fourth Method

This method is sufficient to obtain all solutions for the Lyapounov equation (4-1). Equation (4-1) can be written in the vector form as

$$F\ \bar{x} = \bar{c} \tag{4-42}$$

where $\bar{x}, \bar{c}$ are $n^2 \cdot 1$ elements and F will be $n^2 \cdot n^2$. This method is obviously not suitable for large n.

For A and B and $C \in C^{2 \cdot 2}$ and

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad , \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

104

F can be expressed as

$$F = \begin{bmatrix} a_{11}I - B^T & a_{12}\ I \\ \hline a_{21}I & a_{22}I - B^T \end{bmatrix}$$

Example (4-5): Considering the same problem in example

(4-1)

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & \cdot1 & 1 & 1 \end{bmatrix} \quad , \quad \overline{C} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix}$$

Solving the system

$$F\overline{x} = \overline{c}$$

we obtain the general solution

$$\overline{x} = (ST)\ \overline{c} + N^{4\cdot1}\ Z^{1\cdot1}$$

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

$$= \begin{bmatrix} 1 \\ 1 \\ 1-z \\ z \end{bmatrix} \qquad \text{i.e.,} \quad x = \begin{bmatrix} 1 & 1 \\ 1-z & z \end{bmatrix}$$

consisting condition holds

$$M\bar{c} = \begin{bmatrix} -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix} = 0$$

V Conclusion and Recommendations

Conclusion

An algorithm for computation of various kinds of generalized inverses is established for the matrices over the field of complex numbers. The existence and computation of various kinds of generalized inverses over the ring of polynomials in several variables are studied. Equivalence of a matrix to its Smith form over the ring of polynomials in several variables is studied. A new algorithm for finding the solution of $Ax = b$ over the field of polynomials in several variables is established.

Recommendations

1. Implementation of these algorithms on computer.

2. Study of sufficient and necessary conditions for a matrix over the ring of polynomials in several variables to be equivalent to its Smith form.

3. Explicit solution of Lyaponov and Ricate equation in terms of generalized inverses. Extension of Jones work [15], [16], [17].

4. Applications in the field of control theory. Extention of the work of:

a. Frost and Storey (Contrability and Observability) [11], [12]

b. Das and Ghoshal (Construction of Reduced-order Observes) [7] .

c. Lovass-Nagy, Powers, Al-Nasr [1], [20].

Bibliography

1. Al-Nasr, N., V. Lovass-Nagy, and D. L. Powers, "On Transmission Zeros and Zero Directions of Multivariable Time - Invarient Linear Systems With Input-Derivative Control." Int. J. Contr., Vol. 33, pp. 859 - 870, 1981.

2. Barnett, S., Introduction to Mathematical Control Theory. London: Oxform University Press, 1975.

3. Bose, N.K. and S.K. Mitra, "Generalized Inverse of Polynomial Matrices." IEEE Trans. Automatic Contr., Vol AC-23, pp. 491 - 493, 1978.

4. Ben-Israel, A. Ben and T.N.E. Greville, Generalized Inverses, Theory and Applications. New York: Wiley, 1974.

5. Browne, E.T. Introduction to the Theory of Determinants and Matrices. Richmond, Virginia: The William Byrd Press., 1958.

6. Buslowicz, M..."Inversion of Polynomial Matrices." Int. J. Contr., Vol. 33, pp 977 - 984, 1980.

7. Das, G. and T.K. Ghoshal "Reduced-order Observer Construction by Generalized Matrix Inverse." Int. J. Contr., Vol. 33, pp. 371 - 378, 1981.

8. Frost, M.G. and C. Storey "Equivalence of a Matrix Over R s,z with its Smith Form." Int. J. Contr., Vol. 28, pp. 665 - 671, 1978.

9. Frost, M. G. and C. Storey "Equivalence of Matrices Over R s,z : A Counter-example." Int. J. Contr, Vol. 34, pp. 1225 - 1226, 1981.

10. Frost, M. G. and C. Storey "Transformations of Strict System Equivalence between Polynomial System Matrices Over R s,z ." Int. J. Contr., Vol. 30, pp. 917 - 926, 1979.

11. Frost, M. G. and C. Storey "Further Remarks on the Controllability of Linear Constant Delay-Differential Systems." Int. J. Contr., Vol. 30, pp. 863 - 870, 1979.

12. Frost, M. G. "Controllability, Observability and the Transfer Function Matrix for a Delay-Differential System." Int. J. Contr., Vol. 35, pp. 175 - 182, 1982.

13. Greville, T.N.E. "Solutions of the Matrix Equation XAX = X, and Relations Between Oblique and Orthogonal Projectors." *SIAM J. Appl. Math.*, Vol. 26, No. 4, June 1974.

14. Inouye, Y. "An Algorithm for Inverting Polynomial Matrices." *Int J. Contr.*, Vol. 30, pp. 989 - 999, 1979.

15. Jones, J., Jr. "Solution of Certain Matrix Equations." *Proceedings of the American Mathematical Society*, 31:333 - 339 (1972).

16. Jones, J. Jr., J. Leuthauser, R. Gressang. "Solutions of the Algebraic Riccati Matrix Equation." *Notes of the American Mathematical Society*, 23:17 - 49 (1971).

17. Jones, John, Jr., and Charles Lew. "Solutions of the Lyapunov Matrix Equation $BX-XA=C$." *IEEE Trans. Automatic Contr.*, Vol. AC-27, pp. 464 - 466, 1982.

18. Jones, John, Jr., Direct communication.

19. Lee, E.B. and S.H. Zak, "Smith Form Over $R A_1, Z_2$." *IEEE Trans. Automatic Contr.*, Vol. AC-28, pp. 115 - 118, 1983.

20. Lovass-Nagy, V., R.J. Miller, and D.L. Powers. "Introduction to the Applications of the Simplest Matrix-generalized Inverse in System Science". *IEEE Trans. Circuits Syst.*, Vol. CAS-25, pp. 766 - 771, 1978.

21. Miller, R.J. and R. Mukundan. "On Designing Reduced-order Observers for Linear Time-invariant Systems Subject to Unknown Inputs." *Int. J. Contr.*, Vol. 35, pp. 183 - 188, 1982.

22. Moore

23. Morse, A.S. "Ring Models for Delay-Differential Systems." *Automatica*, Vol. 12, pp. 529 - 531, 1976.

24. Moris, G.L. and P.L. O'Dell. "A Characterization for Generalized Inverses of Matrices." *SIAM Review* Vol. 10, No. 2, pp. 208 - 211, 1968.

25. Penrose

26. Rao, R. and S.K. Mitra, <u>Generalized Inverses of Matrices and Its Applications</u>. New York : Wiley, 1976.

27. Sontag, E.D. "On Generalized Inverses of Polynomial and Other Matrices." <u>IEEE Trans. Automatic Contr.,</u> Vol. AC-25, pp. 514 - 517, 1980.

28. Strang, G. <u>Linear Algebra and Its Applications</u>. New York:Academic Press, 1976.

Appendix A

Basic Applications of Generalized Inverses

Solution of linear equation $Ax=y$:

Theorem (A-1):

A necessary and sufficient condition that $Ax=y$ is consistent is that

$$AA_1y = y \qquad (A-1)$$

The general solution of the consistent equation is

$$x = A_1y + (I - A_1A)Z \qquad (A-2)$$

where Z is an arbitrary vector.

Proof:

Sufficiency: if (A-1) is true, then A_1y is a solution.

Necessity: Suppose that $Ax = y$ is consistent, then there exists w such that

$$Aw = y$$

$$AA_1(Aw) = y$$

$$AA_1 y = y .$$

To complete the proof it is sufficient to prove that (A-2) is a solution for $Ax=y$. Substituting (A-2) into the equation $Ax=y$ we have

$$Ax = A(A_1 y + (I-A_1 A)Z)$$

$$= AA_1 y + AZ - AA_1 AZ$$

$$= y + ZA = AZ$$

$$= y$$

To prove that any solution x can be derived from (A-2), we can choose Z as follows:

$$Z = x - Gy$$

$$x = Gy + (I-GA)Z = Gy + (I-GA)(x-Gy)$$

$$= Gy + x - Gy - GAx + GAGy = x - GAx + GAGAx$$

$$= x - GAx + GAx = x$$

Theorem (A-2):

The necessary and sufficient condition that the equation $AXB=C$ has a solution is that

$$AA_1 CB_1 B = C \quad , \tag{A-3}$$

A-2

in which case the general solution is

$$x = A_1CB_1 + Z - A_1AZBB_1 \qquad (2-4) \qquad\qquad (A-4)$$

where Z is an arbitrary matrix.

Proof:

Sufficiency is trivial since A_1CB_1 is a solution.
Necessity proof: if the equation is consistent, then there exists X such that

$$AXB + C$$

$$AA_1(AXB)B_1B = C$$

$$AA_1CB_1B = C \quad .$$

Substituting X given by (A-4) in AXB, we have $A(A_1CB_1+Z-A_1AZBB_1)B = C+AZB-AZB=C$. Any solution of $AXB=C$ is obtainable through (A-4) by a suitable choice of Z. For example, X can be obtained if we put $Z = X-A_1CB_1$

$$\text{Solution} = A_1CB_1 + (x-A_1CB) - A_1A(x-A_1CB_1)BB_1$$

$$= x - A_1AxBB_1 + A_1 \; A \; A_1 \; C \; B_1 \; B \; B_1$$

$$= x - A_1 \; A \; x \; B \; B_1 + A_1C \; B_1$$

$$= x - A_1 C B_1 + A_1 CB_1 = x$$

Theorem (A-3)

Let $A(m \cdot n)$, $C(m \cdot p)$, $B(p \cdot g)$, $D(n \cdot g)$ be given matrices. A necessary and sufficient condition for the consistent equations $AX=C$, $XB=D$ to have common solution is that

$$AD = CB ,$$

in which case the general expression for a common solution is

$$x = A_1 C + DB_1 - A_1 ADB_1 + (I - A_1 A) Z (I - BB_1)$$

where Z is arbitrary.

Vita

Lt Col Abel-Monem E. Doma was born in Egypt in 1948. After graduating from high school, he attended military technical college, Cairo, Egypt, from which he received a B.S. degree in electronic engineering in 1971. Subsequent assignments included Egyptian Army Signal Corps officer as electronic engineer. In 1976, he was assigned to be an instructor in the Military Technical Institute, Cairo, Egypt. He received his Diploma Degree in Computer System and Automatic Control from Military Technical College in 1981. He entered the Air Force Institute of Technology in June 1982.

AD-A138274

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release and distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	

4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GCS/MA/83D-1	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION School of Engineering	6b. OFFICE SYMBOL (If applicable) AFIT/ENG	7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State and ZIP Code) Air Force Institute of Technology Wright-Patterson, AFB, Ohio 45433	7b. ADDRESS (City, State and ZIP Code)

8a. NAME OF FUNDING/SPONSORING ORGANIZATION Aerospace Medical Research	8b. OFFICE SYMBOL (If applicable) AM RL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State and ZIP Code) AMRL, AFB, OHIO 45433	10. SOURCE OF FUNDING NOS.			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE (Include Security Classification)
See Box 19 (unclassified)

12. PERSONAL AUTHOR(S) Doma, Abdel Monem E.

13a. TYPE OF REPORT MS Thesis	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) 1983 December 16	15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

Approved for public release: IAW AFR 190-17.
LYNN E. WOLAVER 1 Feb 84
Dean for Research and Professional Development
Air Force Institute of Technology (ATC)
Wright-Patterson AFB OH

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Linear Algebra, Generalized inverses, Polynomial matrices, Lyapounov equation, Riccati equation, Least-squares solutions, matrices.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
Title: "Generalized Inverses of Matrices and Its Applications" - Theory and computation techniques of the various types of generalized inverses of matrices which have polynomial elements x, y, z..., etc., are presented. A simple algorithm for computation of generalized inverses of a constant matrix is established, and then applied to the case of matrices having polynomial elements in several variables. Reduction of a matrix to its Smith form over the ring of polynomial elements in several variables is presented. A simply algorithm for investigation of the system Ax = b in case of constant and nonconstant rank of A is presented. Application of generalized inverses to solve more general matrix equations such as Lyapunov and Riccati equations is studied.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION

22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Jones J. Jr.	22b. TELEPHONE NUMBER (Include Area Code) 513-255-3098	22c. OFFICE SYMBOL AFIT/ENG

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

DAT
FILM