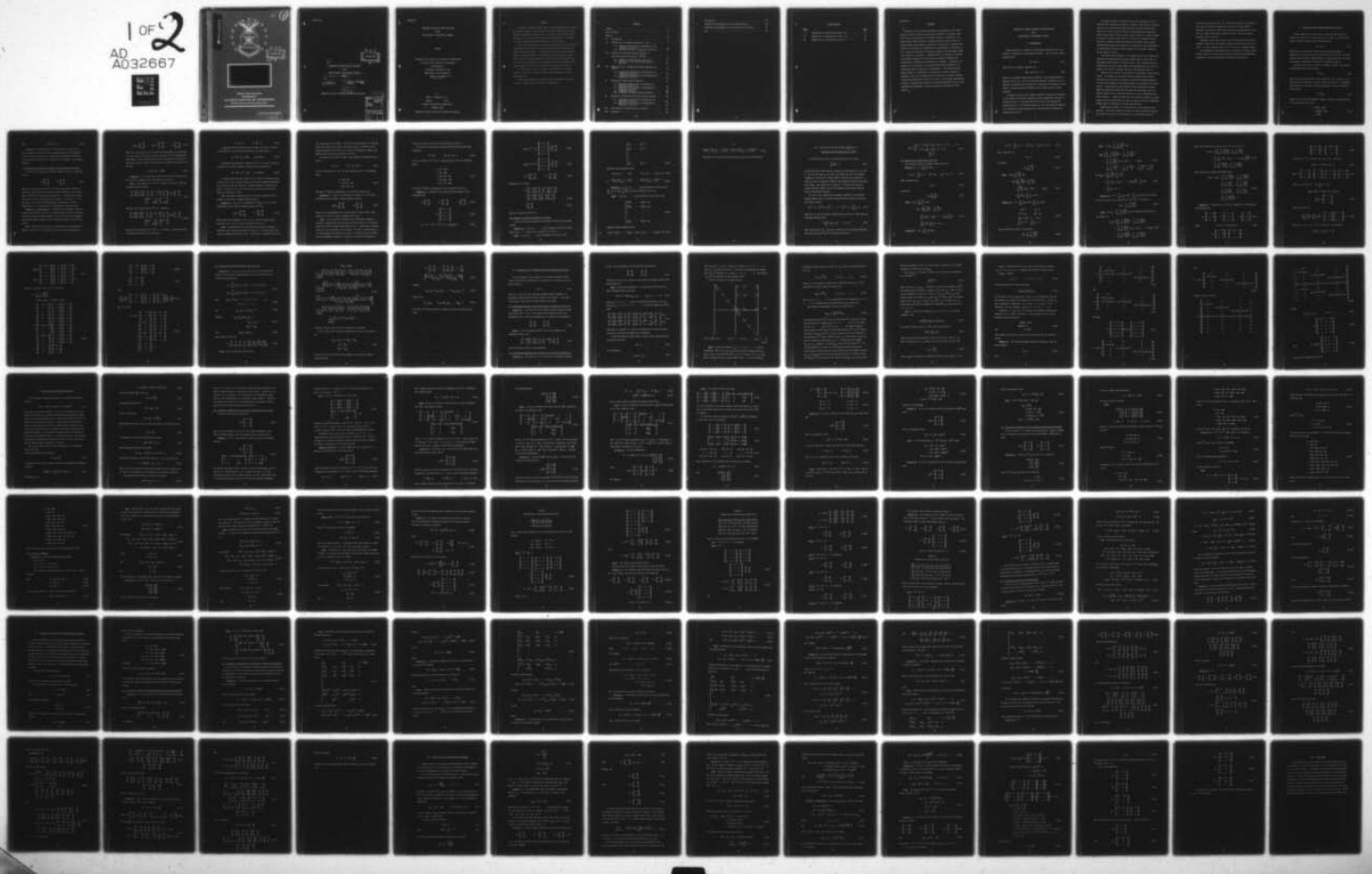
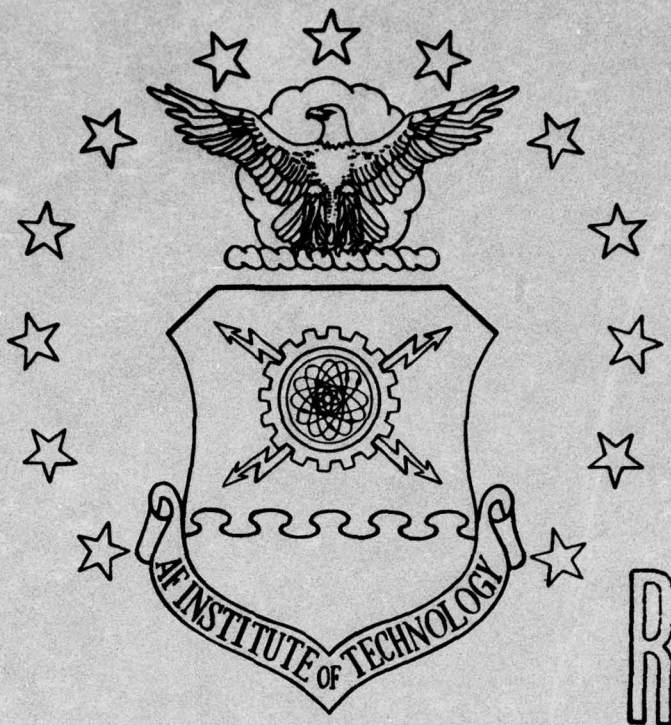
AD-A032 667 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 12/1
SOLUTIONS OF CERTAIN MATRIX EQUATIONS WITH APPLICATIONS TO ENGI--ETC(U)
DEC 76 J L LEUTHAUSER

UNCLASSIFIED GA/MA/76D-1 NL

1 OF 2
AD A032667



AD A032667



DDC
RECEIVED
NOV 29 1976
B

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

GA/MA/76D-1

DDC
RECEIVED
NOV 29 1976
B

(6)

SOLUTIONS OF CERTAIN MATRIX EQUATIONS WITH APPLICATIONS TO ENGINEERING SCIENCES

(9) Masters THESIS

(10) James L. Leuthauser
Captain USAF

(14) GA/MA/76D-1

(11) Dec 76

(12) 104p.

Approved for public release; distribution unlimited

012225

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

GA/MA/76D-1

SOLUTIONS OF CERTAIN MATRIX EQUATIONS WITH APPLICATIONS TO ENGINEERING SCIENCES

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

James L. Leuthauser, B. S.

Captain USAF

Graduate Astronautical Engineering

December 1976

Approved for public release; distribution unlimited

Preface

My studies of optimal 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 that investigation. It would not have been possible without Dr. John Jones, Jr., whose imagination and fertile mind are its true source.

The work was sponsored by the Flight Control Division of the Air Force Flight Dynamics Laboratory, Air Force Wright Aeronautical Laboratories, Wright-Patterson AFB, Ohio. I would particularly like to thank Captain Randall Gressang and Captain Eric Lindberg for their personal support. The results of this paper should be of great value to the United States Air Force in the computation of solutions of many diverse problems which require solutions of the types of equations treated in this thesis.

Ms. Mary Baker, my typist, deserves a big thank you for making beautiful order out of a tangled jungle of matrix equations--a truly monumental task.

My wife and daughter have endured the most throughout this entire AFIT tour. Their patience and love is without bound.

Contents

Preface ii

List of Tables v

Abstract vi

I. Introduction 1

II. Solutions of the Liapunov Equation $AX - XB = C$ 4

2.1. Necessary and Sufficient Conditions for the Solution of Eq (2.1) 5
2.2. Properties of Sets of Solutions of Eq (2.1) . . . 10

III. Solutions of the Matrix Equation $\sum_{i=1}^{k} A_i X B_i = C$ Under More General Conditions on A and B 13

3.1. Solutions of Special Cases of Eq (3.1) 14
3.2. Solutions of the General Linear Case of Eq (3.1) 21

IV. Solutions of $AX = C$ Under More General Conditions on A and C 24

4.1. A Necessary Condition for the Existence of a Solution X of Eq (4.1) 24
4.2. A Sufficient Condition for the Existence of a Solution X of Eq (4.1) 24

V. Solutions of Riccati Matrix Equations 32

5.1. Necessary Conditions for the Existence of Solutions of Eq (5.12) 34
5.2. Sufficient Conditions for the Existence of Solutions of Eq (5.12) 42
5.3. Uniqueness Theorems 46
5.4. A Possible Iterative Solution Technique 56

VI. Solutions of the Generalized Matrix Riccati Equation . . . 59

6.1. A Necessary Condition for the Existence of Solutions of Eq (6.1) 60
6.2. Sufficient Conditions for the Existence of Solutions of Eq (6.1) 61

VII. A New Method for Computing Matrix Inverses 77

VIII. Conclusion 86

Bibliography 87

Supplemental Bibliography for the Liapunov Equation 90

Supplemental Bibliography for the Matrix Riccati Equation 93

Vita 96

List of Tables

Table		Page
I	Combinations of Characteristic Roots of R	51
II	Combinations of Characteristic Roots of R	53
III	Combinations of Characteristic Roots of R	55

Abstract

Solutions of certain matrix equations are presented in this thesis. Results of Mitra are used to extend current methods of obtaining a general form of the solution of the Liapunov equation. Solutions of the general linear matrix equation, including important special cases, are obtained under rather general conditions on the matrices. Necessary and sufficient conditions are established for the existence of a solution of $AX = C$ where the elements of the matrices belong to the polynomial domain of the field of real numbers. Representations of solutions to the algebraic Riccati equation are given along with original techniques for computation of solutions and conditions for uniqueness. Sufficient conditions for the existence of the generalized nonlinear algebraic and differential Riccati equations are established. The classical Newton-Raphson method of finding zeros of functions of a single variable is extended to the case of finding zeros of functions of a matrix variable, including determining the inverse of a matrix. An extensive bibliography includes a comprehensive listing of articles which provide historical and developmental background on the matrix equations considered in this thesis.

SOLUTIONS OF CERTAIN ALGEBRAIC MATRIX EQUATIONS

WITH

APPLICATIONS TO ENGINEERING SCIENCES

I. Introduction

Matrix equations are playing an increasingly important role in the solution of problems in the engineering sciences. Two of the most widely encountered are

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

often called the Liapunov equation, and

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

known as the algebraic matrix Riccati equation. New representations of solutions of Eqs (1.1) and (1.2) along with original techniques for computation of solutions and conditions for uniqueness are presented in this thesis. Included also are treatments of more general forms of these equations.

A growing interest in the Liapunov method of analyzing the stability of both linear and nonlinear systems has meant commensurate interest in solving Eq (1.1), as with Meirovitch (Ref 1:Chap 6) and LaSalle and Lefschetz (Ref 2). The Liapunov equation is also the governing equation in the analysis of beam gridworks with various boundary conditions as studied by Ma (Ref 3).

The digital computer is responsible for the expanding use of the matrix Riccati equation, particularly in modern control theory, which has come into its own during the space program of the last fifteen years. Control of a booster rocket during the launch phase where random wind loads require reaction from an active control system is a typical example of the use of modern control (Ref 4). In fact, as R. Howerton states, "The key to design of optimal controllers and estimators for discrete time-invariant systems with quadratic cost criteria is the determination of the symmetric positive definite solution of the discrete algebraic Riccati equation" (Ref 5).

The matrix Riccati equation is an integral part of the optimal estimation and filtering problem pursued by Kalman (Ref 6), Deyst (Ref 7), Sutherland and Gelb (Ref 8), and many others. Also, orbital trajectory optimization has drawn a large number of investigations such as those of Bryson and Ho (Ref 9), Breakwell (Ref 10), and McReynolds (Ref 11).

Applications of modern control theory are being made in many diverse fields. For example, as the speed of ground transportation systems increases, active control is required to provide the damping necessary for structural integrity and passenger comfort not possible with passive control. Investigation in this field has been accomplished by Bender (Ref 12), Karnopp and Trikha (Ref 13), and Sevin and Pilkey (Ref 14). Recently modern control theory has found use in civil engineering structures analysis. J. Yang has shown that active control systems can insure higher structural reliability and safety in heavy structures by dampening random inputs of earthquakes and high winds (Ref 15).

Modern control theory is not the sole user of the matrix Riccati equation. The positive definite solution of a matrix quadratic equation can replace the spectrum factorization step in the Wiener theory of

filtering and prediction (Ref 16). The Riccati equation also appears in the theory of multiwire transmission lines (Ref 17), and Bellman, Vasudevan, and Ueno have shown that not only the matrix quadratic equation, but higher order matrix equations as well, arise in transport processes (Ref 18; Ref 19).

Throughout this thesis capital letters denote matrices and matrix functions, while lower case letters represent scalars and scalar functions. T is the transpose symbol, and I is the n by n identity matrix.

The supplemental bibliographies are an attempt to include a comprehensive listing of articles which provide historical and developmental background on the matrix equations considered in this thesis.

II. Solutions of the Liapunov Equation AX - XB = C

In this chapter the results of S. K. Mitra (Ref 20) are used to extend the results of W. E. Roth (Ref 21), J. Jones, Jr. (Ref 22), and others to obtain a general form of the solution of

$$AX - XB = C \tag{2.1}$$

This equation is called the Liapunov equation after Aleksander Mikhailovich Liapunov, who lived in Russia from 1857 to 1918. He developed the theory for and created methods of solving problems in the areas of stability, equilibrium, and motion of mechanical systems.

One way for Eq (2.1) to occur is by taking a linear, time-invariant system represented by (Ref 23:184)

$$\dot{\bar{x}} = A\bar{x} \tag{2.2}$$

where $\bar{x}$ is the n state vector of the system and A is a constant n by n matrix which may represent a closed or open loop system. To investigate the stability of the system a potential Liapunov function in quadratic form is used:

$$V = \bar{x}^T P \bar{x} \tag{2.3}$$

where P is an n by n real symmetric matrix. Taking the time derivative of V and using Eq (2.2) gives

$$\begin{aligned} \dot{V} &= \dot{\bar{x}}^T P\bar{x} + \bar{x}^T P\dot{\bar{x}} \\ &= \bar{x}^T A^T P\bar{x} + \bar{x}^T PA\bar{x} \\ &= -\bar{x}^T Q\bar{x} \end{aligned} \tag{2.4}$$

where $$A^T P + PA = -Q \tag{2.5}$$

Employed in this chapter are A^-, defined for any A by $AA^-A = A$, which replaces A^+ used in an earlier representation of the solution of Eq (2.1), and the results of W. E. Roth which established a necessary and sufficient condition for the existence of a solution of the Liapunov equation.

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

If the 2n by 2n matrices R and $\tilde{R}$ are denoted, as in W. E. Roth (Ref 21), by

$$R = \begin{bmatrix} A & C \\ O & B \end{bmatrix}, \qquad \tilde{R} = \begin{bmatrix} A & O \\ O & B \end{bmatrix} \tag{2.6}$$

where A, B, and C are given in Eq (2.1) then a sufficient, though not necessary, condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case the solution of Eq (2.1) not only exists, but is unique. Here $f_A(\lambda) = |A - \lambda I|$ and $f_B(\lambda) = |B - \lambda I|$ are the characteristic polynomials of A and B respectively. W. E. Roth (Ref 21) established the following result:

Theorem 2.1 - Roth (Ref 21). A necessary and sufficient condition that Eq (2.1), where A, B, and C are square matrices of order n with elements in the field of complex numbers, have a solution X with elements in the field of complex numbers, is that the matrices given in Eq (2.6) be similar.

Next, a solution X of Eq (2.1) is obtained under the hypotheses of Theorem 2.1 above. If the following 2n by 2n matrices are denoted by

$$f_A(R) = \begin{bmatrix} 0 & M \\ 0 & U \end{bmatrix}, \qquad f_B(R) = \begin{bmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \tag{2.7}$$

where the n by n matrices U, M, $\hat{M}$, and $\hat{N}$ are polynomials in the matrices A, B, and C given in Eq (2.1) and $f_A(\lambda)$ and $f_B(\lambda)$ are the characteristic polynomials of A and B respectively, then the hypotheses of Theorem 2.1 above can be shown to imply that there exists a solution X of Eq (2.1) of the form

$$X = \hat{N}^{-}\hat{M} - MU^{-} + \hat{N}^{-}\hat{N}MU^{-} \tag{2.8}$$

Theorem 2.2. If the matrices R and $\tilde{R}$ of Eq (2.6) are similar, then Eq (2.1) has a solution X of the form given in Eq (2.8).

Proof. From Theorem 2.1, if X is a solution of Eq (2.1), then the following equation results:

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & M \\ 0 & U \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & M + XU \\ 0 & U \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} f_A(R) \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} f_A(A) & 0 \\ 0 & f_A(B) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & f_A(B) \end{bmatrix} \tag{2.9}$$

Equation (2.9) implies that $M + XU = 0$. Similarly,

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} \hat{N} & -\hat{N}X + \hat{M} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} f_B(R) \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} f_B(A) & 0 \\ 0 & f_B(B) \end{bmatrix} = \begin{bmatrix} f_B(A) & 0 \\ 0 & 0 \end{bmatrix} \tag{2.10}$$

Equation (2.10) implies that $\hat{M} - \hat{N}X = 0$. Therefore, a solution X of Eq (2.1) must satisfy the pair of equations

$$M + XU = 0, \qquad \hat{M} - \hat{N}X = 0 \tag{2.11}$$

A necessary and sufficient condition for $\hat{N}X = \hat{M}$ to have a solution is that $\hat{N}\hat{N}^{-}\hat{M} = \hat{M}$, and the general solution X_1 is then given by

$$X_1 = \hat{N}^{-}\hat{M} + Y_1 - \hat{N}^{-}\hat{N}Y_1 \qquad Y_1 \text{ arbitrary} \tag{2.12}$$

A necessary and sufficient condition for $XU = -M$ to have a solution is that $MU^{-}U = M$, and the general solution X_2 is then given by

$$X_2 = -MU^{-} + Y_2 - Y_2UU^{-} \qquad Y_2 \text{ arbitrary} \tag{2.13}$$

A necessary and sufficient condition for the pair of equations given in Eq (2.11) to have a common solution is that each equation individually has a solution, and that $-\hat{N}M = \hat{M}U$. A common solution X is given by Eq (2.8). A solution $X = X_1 = X_2$ of Eq (2.1) is of the form in Eq (2.8) where $Y_2 = \hat{N}^{-}M$ and $Y_1 = -MU^{-}$.

Other sufficient conditions that Eq (2.1) have a solution X with elements in the field of complex numbers follow.

Theorem 2.3. If $f_\alpha(\lambda)$ is a polynomial of degree $n \geq 1$ in λ with coefficients in the field of complex numbers such that

$$f_\alpha(R) = \begin{bmatrix} V & N \\ 0 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \tag{2.14}$$

where V and N are polynomials in A, B, and C, and V^{-1} exists, then a solution X of $N - VX = 0$ is also a solution of Eq (2.1).

Proof. As previously stated a sufficient, though not necessary, condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case a solution of Eq (2.1)

not only exists, but is unique. But this is also necessary, for the hypothesis here implies that $f_\alpha(B) = 0$, whence $f_\alpha(\lambda)$ is a multiple of the minimum polymonial of B, and that $V = f_\alpha(A)$ is nonsingular, whence $f_\alpha(\lambda)$ cannot vanish for any characteristic root of A.

The matrices $f_\alpha(R)$ and R commute, which implies the following identities:

$$AV = VA, \qquad AN - VC - NB = 0 \tag{2.15}$$

If X is a solution of $N - VX = 0$, then, using Eq (2.15), the following holds:

$$\begin{aligned} 0 &= A(N - VX) \\ &= VC + NB - AVX \\ &= V(C + XB - AX) \end{aligned} \tag{2.16}$$

and since V^{-1} exists by hypothesis, X is a solution of Eq (2.1).

Theorem 2.4. If $f_\beta(\lambda)$ is a polynomial of degree $n \geq 1$ in λ with coefficients in the field of complex numbers such that

$$f_\beta(R) = \begin{bmatrix} 0 & N \\ 0 & M \end{bmatrix}, \qquad R = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \tag{2.17}$$

where N, M are polynomials in A, B, and C, and M^{-1} exists, then a solution X of $N + XM = 0$ is also a solution of Eq (2.1).

Proof. As in the proof of Theorem 2.3, a sufficient, though not necessary, condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case a solution of Eq (2.1) not only exists, but is unique. But this is also necessary, for the hypothesis here implies that $f_\beta(\Lambda) = 0$, whence $f_\beta(\lambda)$ is a multiple of the minimum polynomial of A and that $M = f_\beta(B)$ is nonsingular,

whence $f_\beta(\lambda)$ cannot vanish for any characteristic root of B.

The matrices of $f_\beta(R)$ and R commute which implies the following identities:

$$BM = MB, \qquad AN + CM - NB = 0 \tag{2.18}$$

If X is a solution of $N + XM = 0$, then, using Eq (2.18), the following holds:

$$\begin{aligned} 0 &= -(N + XM)B \\ &= -NB - XMB \\ &= -CM - AN - XMB \\ &= -CM + AXM - XMB \\ &= (-C + AX - XB)M \end{aligned} \tag{2.19}$$

and since M^{-1} exists by hypothesis, X is a solution of Eq (2.1).

Example 2.1. Solving Eq (2.1) with the above technique is now illustrated for

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \tag{2.20}$$

$$R = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{2.21}$$

$$|R - \lambda I| = -(\lambda^2 - \lambda)(\lambda^2 + \lambda) = -f_A(\lambda) f_B(\lambda) \tag{2.22}$$

$$f_A(R) = R^2 - R = \begin{bmatrix} 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & M \\ 0 & U \end{bmatrix} \tag{2.23}$$

$$f_B(R) = R^2 + R = \begin{bmatrix} 2 & 0 & 2 & 2 \\ 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{bmatrix} \tag{2.24}$$

$$\hat{N}^- = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \end{bmatrix}, \qquad U^- = \begin{bmatrix} 0 & 1/2 \\ 0 & 1/2 \end{bmatrix} \tag{2.25}$$

Using Eq (2.8) to find X

$$X = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 0 & -4 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1/2 \\ 0 & 1/2 \end{bmatrix} + \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1/2 \\ 0 & 1/2 \end{bmatrix} \tag{2.26}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \tag{2.27}$$

which is a solution of Eq (2.1).

2.2 Properties of Sets of Solutions of Eq (2.1)

In this section some properties of solutions of Eq (2.1) are established.

Theorem 2.5. If $X_1, X_2, \ldots, X_n$ are solutions of Eq (2.1), then $\frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ is also a solution of Eq (2.1).

Proof. If $X_1, X_2, \ldots, X_n$ are solutions of Eq (2.1), then

$$\begin{cases} AX_1 - X_1B = C \\ AX_2 - X_2B = C \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ AX_n - X_nB = C \end{cases} \tag{2.28}$$

Adding the above equations gives

$$A(X_1 + X_2 + \cdot \cdot \cdot + X_n) - (X_1 + X_2 + \cdot \cdot \cdot X_n)B = nC \tag{2.29}$$

or
$$\frac{A(X_1 + X_2 + \cdot \cdot \cdot + X_n)}{n} - \frac{(X_1 + X_2 + \cdot \cdot \cdot X_n)B}{n} = C \tag{2.30}$$

Theorem 2.6. If $X_1, X_2, \ldots, X_n$ are solutions of Eq (2.1) and $\sum_{i=1}^{n} \alpha_i = m$, then $\sum_{i=1}^{n} \frac{\alpha_i}{m} X_i$ is also a solution of Eq (2.1).

Proof. If $X_1, X_2, \ldots, X_n$ are solutions of Eq (2.1), then

$$\begin{cases} \alpha_1 AX_1 - \alpha_1 X_1 B = \alpha_1 C \\ \alpha_2 AX_2 - \alpha_2 X_2 B = \alpha_2 C \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \alpha_n AX_n - \alpha_n X_n B = \alpha_n C \end{cases} \tag{2.31}$$

Adding the above equations gives

$$A(\alpha_1 X_1 + \alpha_2 X_2 + \cdot \cdot \cdot + \alpha_n X_n) - (\alpha_1 X_1 + \alpha_2 X_2 + \cdot \cdot \cdot + \alpha_n X_n)B = mC \tag{2.32}$$

or

$$\frac{A(\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n)}{m} - \frac{(\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n)B}{m} = C \qquad (2.33)$$

Therefore, Eq (2.33) is the solution form stated in the theorem above.

III. Solutions of the Matrix Equation $\sum_{i=1}^{k} A_i X B_i = C$

Under More General Conditions on A and B

In this chapter solutions of matrix equations of the form

$$\sum_{i=1}^{k} A_i X B_i = C \tag{3.1}$$

are obtained under rather general conditions on the matrices A_i, B_i, and C. For the case where A_i, B_i, and C are square matrices with elements belonging to the field of complex numbers, solutions are obtained in terms of C and the principal idempotent and nilpotent matrices associated with $\{A_i\}_{i=1}^{k}$, and $\{B_i\}_{i=1}^{k}$ (Ref 24:Chap 16). Solutions are found for certain important special cases of the general linear matrix equation such as the Liapunov equation.

If A is a n by n matrix having elements belonging to the field of complex numbers and if the reduced characteristic function $\phi(\lambda)$ factors into linear factors,

$$\phi(\lambda) = (\lambda - a_1)^{\nu_1}(\lambda - a_2)^{\nu_2} \ldots (\lambda - a_s)^{\nu_s}, \quad s \geq 2, \quad \sum_{i=1}^{m} \nu_i \leq n \tag{3.2}$$

where the $\{a_i\}$ are the distinct characteristic roots of A, then A has the following representation:

$$A = \sum_{j=1}^{m} (a_j A_j + \bar{A}_j), \qquad 2 \leq m \leq n \tag{3.3}$$

where the matrices $\{A_j\}$, $\{\bar{A}_j\}$ form a complete set of principal idempotent and nilpotent matrices with the following properties:

$$A_jA_k = \begin{cases} A_j,\ j = k;\ A_j\bar{A}_j = \bar{A}_jA_j = \bar{A}_j; & j = 1, 2, \ldots, m \\ 0\ ,\ j \neq k;\ A_j\bar{A}_k = \bar{A}_jA_k = 0\ ; & j \neq k \end{cases} \tag{3.4}$$

$$\sum_{j=1}^{m} A_j = I$$

3.1 Solutions of Special Cases of Eq (3.1)

The following results are special cases of Eq (3.1):

Theorem 3.1. If $A = \sum_{i=1}^{n} a_iE_i$ where

$$a_i \neq 0; \quad \sum_{i=1}^{n} E_i = I; \quad E_i^2 = E_i; \quad E_iE_j = 0,\ i \neq j \tag{3.5}$$

then a solution X of

$$AX = C \tag{3.6}$$

is given by

$$X_A = \sum_{i=1}^{n} \frac{E_i}{a_i} C \tag{3.7}$$

Proof. If $X_A = \sum_{i=1}^{n} \frac{E_i}{a_i} C$ then

$$AX_A = A\left(\sum_{i=1}^{n} \frac{E_i}{a_i}\right)C = \sum_{i=1}^{n} \frac{AE_i}{a_i} C$$

$$= \sum_{i=1}^{n} (a_1E_1 + a_2E_2 + \cdots + a_nE_n) \frac{E_i}{a_i} C \tag{3.8}$$

$$= \sum_{i=1}^{n} E_i^2C = \sum_{i=1}^{n} E_iC = IC = C$$

Theorem 3.2. If $B = \sum_{j=1}^{n} b_jF_j$ where

$$b_j \neq 0; \quad \sum_{j=1}^{n} F_j = I; \quad F_j^2 = F_j; \quad F_iF_j = 0, \ i \neq j \tag{3.9}$$

then a solution X of

$$XB = D \tag{3.10}$$

is given by

$$X_B = \sum_{j=1}^{n} \frac{DF_j}{b_j} \tag{3.11}$$

Proof. If $X_B = \sum_{j=1}^{n} \frac{DF_j}{b_j}$ then

$$X_B B = \left(\sum_{j=1}^{n} \frac{DF_j}{b_j} \right) B = \sum_{j=1}^{n} \frac{DF_j}{b_j} B$$

$$= \sum_{j=1}^{n} \frac{DF_j}{b_j} (b_1F_1 + b_2F_2 + \cdots + b_nF_n) \tag{3.12}$$

$$= \sum_{j=1}^{n} DF_j^2 = \sum_{j=1}^{n} DF_j = DI = D$$

Theorem 3.3. If $A = \sum_{i=1}^{n} a_iE_i$ and $B = \sum_{j=1}^{n} b_jF_j$ where

$$\begin{aligned} &a_i \neq 0, && b_j \neq 0 \\ &E_i^2 = E_i, && F_j^2 = F_j \\ &E_iE_j = F_iF_j = 0, && i \neq j \\ &\sum_{i=1}^{n} E_i = I, && \sum_{j=1}^{n} F_j = I \end{aligned} \tag{3.13}$$

then a solution X of AXB = G is given by

$$X_{AB} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_iGF_j}{a_ib_j} \tag{3.14}$$

Proof. If $X_{AB} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j}{a_i b_j}$ then

$$X_{AB}B = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j}{a_i b_j} \right) B = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j}{a_i b_j} B$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j}{a_i b_j} (b_1 F_1 + b_2 F_2 + \cdots + b_n F_n) \tag{3.15}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j^2}{a_i} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i G F_j}{a_i} = \sum_{i=1}^{n} \frac{E_i G I}{a_i} = \sum_{i=1}^{n} \frac{E_i G}{a_i}$$

and $AX_{AB}B = A\left(\sum_{i=1}^{n} \frac{E_i G}{a_i} \right) = \sum_{i=1}^{n} A \frac{E_i G}{a_i}$

$$= \sum_{i=1}^{n} (a_1 E_1 + a_2 E_2 + \cdots + a_n E_n) \frac{E_i G}{a_i} = \sum_{i=1}^{n} E_i^2 G \tag{3.16}$$

$$= \sum_{i=1}^{n} E_i G = IG = G$$

Theorem 3.4. If $A = \sum_{i=1}^{n} a_i E_i$ and $B = \sum_{j=1}^{n} b_j F_j$ where the relations in Eq (3.12) hold, then a solution X of $AX - XB = C$ is given by

$$X_{A-B} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j}{a_i - b_j} \tag{3.17}$$

Proof. If $X_{A-B} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j}{a_i - b_j}$ then by multiplying both sides on the right by B

$$X_{A-B}B = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j}{a_i - b_j} \right) B = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j}{a_i - b_j} B$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j}{a_i - b_j} (b_1 F_1 + b_2 F_2 + \cdots + b_n F_n) \tag{3.18}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j^2 b_j}{a_i - b_j} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{E_i C F_j b_j}{a_i - b_j}$$

Now by multiplying both sides of Eq (3.17) on the left by A

$$AX_{A-B} = A\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{E_i CF_j}{a_i - b_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} A\,\frac{E_i CF_j}{a_i - b_j}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}(a_1E_1 + a_2E_2 + \cdots + a_nE_n)\,\frac{E_i CF_j}{a_i - b_j} \qquad (3.19)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i E_i^2 CF_j}{a_i - b_j} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i E_i CF_j}{a_i - b_j}$$

Then forming the following differences gives

$$\begin{aligned} AX_{A-B} - X_{A-B}B &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i E_i CF_j}{a_i - b_j} - \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{E_i CF_j b_j}{a_i - b_j} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i ICI}{a_i - b_j} - \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{ICIb_j}{a_i - b_j} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i C}{a_i - b_j} - \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{Cb_j}{a_i - b_j} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(a_i - b_j)C}{a_i - b_j} = C \end{aligned} \qquad (3.20)$$

Example 3.1. Equation (2.1) is used to demonstrate the techniques just presented for

$$A = \begin{bmatrix} 3 & -7 & -20 \\ 0 & -5 & -14 \\ 0 & 3 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \qquad (3.21)$$

where

$$|A - \lambda I| = (\lambda - 1)(\lambda - 2)(\lambda - 3) \qquad (3.22)$$

$$A = \begin{bmatrix} 3 & -7 & -20 \\ 0 & -5 & -14 \\ 0 & 3 & 8 \end{bmatrix} = 1\begin{bmatrix} 0 & -11/2 & -11 \\ 0 & 7 & 14 \\ 0 & -3 & -6 \end{bmatrix} +$$

$$2\begin{bmatrix} 0 & 18 & 42 \\ 0 & -6 & -14 \\ 0 & 3 & 7 \end{bmatrix} + 3\begin{bmatrix} 1 & -25/2 & -31 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.23}$$

with $\bar{A}_1 = \bar{A}_2 = \bar{A}_3 = 0$ and $A = 1A_1 + 2A_2 + 3A_3$. Now for B,

$$|B-\lambda I| = (\lambda-2)^2(\lambda-1),\ \bar{B}_2 = 0 \tag{3.24}$$

$$\text{and}\quad B = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix} = 2\begin{bmatrix} 3 & 3 & 0 \\ -2 & -2 & 0 \\ -2 & -3 & 1 \end{bmatrix} + \begin{bmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{bmatrix} + 1\begin{bmatrix} -2 & -3 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix} \tag{3.25}$$

with $B = 2B_1 + \bar{B}_1 + 1B_2$. Now $\sum_{j=1}^{3} \bar{A}_j = 0$ and

$$A_2\bar{A}_1 = \bar{A}_1A_2 = 0,\ A_1\bar{A}_1 = \bar{A}_1 = \bar{A}_1A_1 \text{ and}$$

$$\sum_{j=1}^{2} \bar{B}_j = \begin{bmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{bmatrix} \tag{3.26}$$

Next it is seen that

$$C\left(\sum_{j=1}^{2} \bar{B}_j\right) = \left(\sum_{k=1}^{3} \bar{A}_k\right)C = \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix}\begin{bmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{bmatrix} = 0 \tag{3.27}$$

and $a_1 = 1$, $a_2 = 2$, $a_3 = 3$, $b_1 = 2$, and $b_2 = 1$. Therefore,

$$(a_1-b_2) = (a_2-b_1) = 0, \text{ and}$$

$$A_1CB_2 = \begin{bmatrix} 0 & -11/2 & -11 \\ 0 & 7 & 14 \\ 0 & -3 & -6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} -2 & -3 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix} = 0$$

$$A_2CB_1 = \begin{bmatrix} 0 & 18 & 42 \\ 0 & -6 & 14 \\ 0 & 3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ -2 & -2 & 0 \\ -2 & -3 & 1 \end{bmatrix} = 0 \qquad (3.28)$$

Finally, a solution X of AX - XB = C is given by

$$X = \sum_{k=1}^{3} \sum_{j=1}^{2} \frac{(A_k C B_j)}{(a_k - b_j)}$$

$$= -A_1CB_1 + A_2CB_2 + 1/2\ A_3CB_2 + A_3CB_1$$

$$= \begin{bmatrix} 0 & 11/2 & 11 \\ 0 & -7 & -14 \\ 0 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ -2 & -2 & 0 \\ -2 & -3 & 1 \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 18 & 42 \\ 0 & -6 & -14 \\ 0 & 3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} -2 & -3 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} 2/4 & -25/4 & -31/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} -2 & -3 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} 1 & -25/2 & -31 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ -2 & -2 & 0 \\ -2 & -3 & 1 \end{bmatrix} \qquad (3.29)$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 12 & 18 & 0 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} -5 & -15/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.30}$$

$$= \begin{bmatrix} 6 & 17/2 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \tag{3.31}$$

Also,

$$X\left(\sum_{j=1}^{3} \overline{B}_j\right) = \begin{bmatrix} -6 & -17/2 & -1 \\ 4 & 6 & 0 \\ -2 & -3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 6 & -3 \\ -2 & -4 & 2 \\ -1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \left(\sum_{k=1}^{2} \overline{A}_k\right) X \tag{3.32}$$

and, as a check,

$$\begin{aligned} AX - XB = & \begin{bmatrix} 3 & -7 & -20 \\ 0 & -5 & -14 \\ 0 & 3 & 8 \end{bmatrix} \begin{bmatrix} 6 & 17/2 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} - \\ & \begin{bmatrix} 6 & 17/2 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 \\ 0 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix} \\ = & \begin{bmatrix} 6 & 15/2 & 3 \\ -8 & -12 & 0 \\ 4 & 6 & 0 \end{bmatrix} - \begin{bmatrix} 5 & 13/2 & 2 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} \\ = & \begin{bmatrix} 1 & 1 & 1 \\ -4 & -6 & 0 \\ 2 & 3 & 0 \end{bmatrix} = C \end{aligned} \tag{3.33}$$

3.2. Solution of the General Linear Case of Eq (3.1)

Theorem 3.5. If A_1, A_2, B_1, B_2, and C are n by n matrices with elements in the field of complex numbers, and if A_i and B_i have the following representations:

$$A_i = \sum_{j=1}^{n_i} (a_{ij} A_{ij} + \overline{A}_{ij}), \quad i = 1, 2; \ 1 \leq n_i \leq n$$

$$B_i = \sum_{k=1}^{n_i} (b_{ik} B_{ik} + \overline{B}_{ik}), \quad i = 1, 2; \ 1 \leq n_i \leq n \tag{3.34}$$

with $$\overline{A}_{ij} C = C\overline{B}_{ik} = 0, \quad i = 1, 2; \ j = 1, 2, \ldots, n_i$$
$$k = 1, 2, \ldots, n_i \tag{3.35}$$

and $$a_{1j_1} b_{1k_1} + a_{2j_2} b_{2k_2} = 0 \tag{3.36}$$

whenever $$A_{ij_i} A_{2j_2} CB_{1k_1} B_{2k_2} = 0$$

$$CA_{ij} = A_{ij} C \tag{3.37}$$

$$B_{ik} C = CB_{ik}$$

then $$A_1 X B_1 + A_2 X B_2 = C \tag{3.38}$$

has a solution X given by

$$X = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \frac{\left(A_{1j_1} A_{2j_2} CB_{1k_1} B_{2k_2}\right)}{\left(a_{1j_1} b_{1k_1} + a_{2j_2} b_{2k_2}\right)} \tag{3.39}$$

Proof. If X is given by Eq (3.39), then

$$A_1XB_1 + A_2XB_2$$

$$= \sum_{\substack{1\le j_i\le n_i \\ 1\le k_i\le n_i}} \left\{ \frac{A_1\, A_{1j_1}\, A_{2j_2}\, CB_{1k_1}\, B_{2k_2}\, B_1}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} + \frac{A_2\, A_{1j_1}\, A_{2j_2}\, CB_{1k_1}\, B_{2k_2}\, B_2}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} \right\}$$

$$= \sum_{\substack{1\le j_i\le n_i \\ 1< k_i\le n_i}} \left\{ \frac{\left(\sum_{\lambda=1}^{n_1} a_{1\lambda} A_{1\lambda} + \bar{A}_{1\lambda}\right)\left(A_{1j_1}\, A_{2j_2}\, CB_{1k_1}\, B_{2k_2}\right)\left(\sum_{\lambda=1}^{n_1} b_{1\lambda_1} B_{1\lambda_1} + \bar{B}_{1\lambda}\right)}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} + \right. \tag{3.40}$$

$$\left. \frac{\left(\sum_{\mu=1}^{n_2} a_{2\mu}\, A_{2\mu} + \bar{A}_{2\mu}\right)\left(A_{1j_1}\, A_{2j_2}\, CB_{1k_1}\, B_{1k_1}\, B_{2k_2}\right)\left(\sum_{\mu=1}^{n_2} b_{2\mu}\, B_{2\mu} + \bar{B}_{2\mu}\right)}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} \right\}$$

$$= \sum_{\substack{1\le j_i\le n_i \\ 1\le k_i\le n_i}} \left\{ \frac{a_{1j_1} b_{1k_1} A_{1j_1} A_{2j_2} CB_{1k_1} B_{2k_2}}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} + \frac{a_{2j_2} b_{2k_2} A_{1j_1} A_{2j_2} CB_{1k_1} B_{2k_2}}{(a_{1j_1}\, b_{1k_1} + a_{2j_2}\, b_{2k_2})} \right\}$$

$$\sum_{\substack{1\le j_i\le n_i \\ 1\le k_i\le n_i}} \left\{ A_{1j_1}\, A_{2j_2}\, CB_{1k_1}\, B_{2k_2} \right\} = C$$

Therefore, X given by Eq (3.39) is a solution of Eq (3.38).

If $\{A_i\}$, $\{B_i\}$, and C have the representations listed in section 3.1 and

$$C\bar{A}_{ij} = \bar{A}_{ij}C = 0 = \bar{B}_{ik}C = C\bar{B}_{ik}$$

$$A_{ij}C = CA_{ij} \tag{3.41}$$

$$B_{ik}C = CB_{ik}$$

then Eq (3.1) has a solution X with elements in the field of complex numbers given by

$$X = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_h=1}^{n_h} \sum_{j_1=1}^{n_1} \sum_{j_2=2}^{n_2} \cdots \sum_{j_h=1}^{n_h}$$

$$\left\{ \frac{A_{1i_1} A_{2i_2} \cdots A_{hi_h} CB_{1k_1} B_{2k_2} \cdots B_{hk_h}}{a_{1j_1} b_{1k_1} + \cdots + a_{hi_h} b_{hk_h}} \right\} \tag{3.42}$$

whenever

$$a_{1j_1} b_{1k_1} + a_{2j_2} b_{2k_2} + \cdots + a_{hj_h} b_{hk_h} = 0 \tag{3.43}$$

implies that

$$A_{1i_1} A_{2i_2} \cdots A_{hi_h} CB_{1k_1} B_{2k_2} \cdots B_{hk_h} = 0 \tag{3.44}$$

The proof of the above theorem is similar to those in the earlier sections.

IV. Solutions of AX = C Under More General Conditions on A and C

The main purpose of this chapter is to establish necessary conditions and sufficient conditions for the existence of a solution X of the matrix equation

$$AX = C \tag{4.1}$$

where A, X, and C are n by n matrices having elements belonging to a polynomial domain $\mathfrak{I}(x)$ where $\mathfrak{I}$ is the field of real numbers. Such equations arise in the work of B. Porter (Ref 25) and others.

4.1 A Necessary Condition for the Existence of a Solution X of Eq (4.1)

Theorem 4.1. *If A and C are n by n matrices having elements belonging to* $\mathfrak{I}(x)$, and if X is a n by n matrix with elements in $\mathfrak{I}(x)$ which satisfies Eq (4.1), then the following 2n by 2n matrices are equivalent:

$$\begin{bmatrix} A & C \\ 0 & I \end{bmatrix}, \qquad \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \tag{4.2}$$

Proof. If X is a solution of Eq (4.1) with the elements of X belonging to $\mathfrak{I}(x)$, then

$$\begin{bmatrix} I & -AX \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & C-AX \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \tag{4.3}$$

and the matrices of Eq (4.2) are equivalent.

4.2 A Sufficient Condition for the Existence of a Solution X of Eq (4.1)

Theorem 4.2. If A and C are n by n matrices with elements belonging

to $\mathcal{J}(x)$, and the following 2n by 2n matrices are equivalent

$$\begin{bmatrix} A & C \\ 0 & A \end{bmatrix}, \qquad \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \tag{4.4}$$

then there exists a n by n matrix X with elements belonging to $\mathcal{J}(x)$ which satisfies Eq (4.1).

Proof. There exist nonsingular n by n matrices P and Q with elements belonging to $\mathcal{J}(x)$ such that

$$PAQ = A' = \text{diag}\,\{a_{11}, a_{22}, \ldots, a_{\alpha\alpha}, 0, \ldots, 0\} \tag{4.5}$$

where a_{ii}, $i = 1, 2, \ldots \alpha$ are the invariant factors of A, i.e., a_{ii} divides a_{jj}, $i \leq j$, $i, j = 1, 2, \ldots, \alpha$. If the following definitions exist

$$\begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} A & C \\ 0 & A \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} PA & PC \\ 0 & PA \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} PAQ & PCQ \\ 0 & PAQ \end{bmatrix} = \begin{bmatrix} A' & C' \\ 0 & A' \end{bmatrix} \triangleq M$$

$$\begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} PA & 0 \\ 0 & PA \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} PAQ & 0 \\ 0 & PAQ \end{bmatrix} = \begin{bmatrix} A' & 0 \\ 0 & A' \end{bmatrix} \triangleq N \tag{4.6}$$

where $PAQ = A'$ and $PCQ = C'$, then M is equivalent to N, since the matrices given in Eq (4.4) are assumed to be equivalent.

It will be shown that there exists a matrix U with elements belonging to $\mathcal{J}(x)$ such that

$$A'U = C' \tag{4.7}$$

or, by elements,

$$a'_{ii}u_{ij} = c'_{ij} \tag{4.8}$$

where $a'_{ii} \in A'$, $c'_{ij} \in C'$, and $u_{ij} \in U$, where $i, j = 1, 2, \ldots, n$, and $a'_{ii} = a_{ii}$ given in Eq (4.5). This will be accomplished by showing that c'_{ij} is a multiple of a'_{ii} for $i, j = 1, 2, \ldots, n$. The elements c'_{ij} will be considered in four separate cases.

The 2n by 2n matrix M is given below by

$$M = \left[\begin{array}{ccccc|ccccc}
a'_{11} & & & & & c'_{11} & \cdots & c'_{1j} & \cdots & c'_{1n} \\
& a'_{22} & & & & \vdots & & \vdots & & \vdots \\
& & \ddots & & & & & & & \\
& & & a'_{ii} & & c'_{i1} & \cdots & c'_{ij} & \cdots & c'_{in} \\
& & & & \ddots & \vdots & & \vdots & & \vdots \\
& & & & a'_{\alpha\alpha} & & & & & \\
& & & & & \vdots & & \vdots & & \vdots \\
& & & & 0 & & & & & \\
& & & & 0 & c'_{n1} & & c'_{ni} & & c'_{nn} \\
\hline
& & & & & a'_{11} & & & & \\
& & & & & & a'_{22} & & & \\
& & & & & & & \ddots & & \\
& & & & & & & & a'_{jj} & \\
& & & & & & & & & \ddots \\
& & & & & & & & & a'_{\alpha\alpha} \\
& & & & & & & & & \ddots \\
& & & & & & & & & 0 \\
& & & & & & & & & 0
\end{array}\right] \qquad (4.9)$$

and $N = \text{diag}\{a'_{11}, \ldots, a'_{\alpha\alpha}, 0 \ldots 0, a'_{11}, \ldots, a_{\alpha\alpha}, 0 \ldots 0\}$ (4.10)

Case 1. First the elements c'_{ij}, $i, j = 1, 2, \ldots, \alpha$ will be considered. Now if a'_{ii} and a'_{jj} have the greatest common factor g_{ij} ($g_{ij} = a'_{ii}$), it can be shown that c'_{ij} is its multiple (of $a'_{ii} = g_{ij}$) and as a result Eq (4.8) is again valid in $\nabla(x)$. If g is any factor

irreducible in $\nabla(x)$ which is a factor of $a'_{\alpha\alpha}$, then the invariant factors of A are

$$a'_{ii} = g^{r_i} \bar{a}_{ii}, \qquad r_1 \leq r_2 \leq \ldots \leq r_\alpha \tag{4.11}$$

where $\bar{a}_{ii}$ is a polynomial in $\nabla(x)$ and is relatively prime to g. Consequently the invariant factors of M (of N) are

$$m_{kk} = g^{t_k} m'_{kk}, \qquad t_1 \leq t_2 \leq \ldots \leq t \tag{4.12}$$

where t_k, k = 1, 2, . . . 2α, is a permutation of the exponents r_i, i = 1, 2, . . . 2α, in nondescending order and where m'_{kk} is prime to g.

The (i+j-1)th determinantal divisor of M (and of N) is

$$M_{i+j-1} = \Pi_{k=1}^{i+j-1} g^{t_k} m'_{kk} \tag{4.13}$$

It contains the factor g^{r_i} of a'_{ii} or the factor g^{r_j} of a'_{jj} for in forming the sequence of integers $\{t_k, k = 1, 2, \ldots, 2i-1\} = \{t\}$, either all $\{r_1, r_2, \ldots, r_i\}$ or all $\{r_1, r_2, \ldots, r_j\}$ must be used in order to get its i+j-1 terms. On the other hand both g^{r_i} and g^{r_j} cannot be factors of M_{i+j-1} for both r_i and r_j cannot occur in the sequence {t} which has only i+j-1 terms. Neither r_i nor r_j is zero since a'_{ii} and a'_{jj} are not relatively prime unless $a'_{ii} = 1$. Consequently the minimum $\{r_i, r_j\} = t_{ij}$ belongs to {t} and M_{i+j-1} must have the factor $g^{t_{ij}}$. Now it is assumed that M' and N' are the matrices obtained from M and N respectively by deleting their ith and (n+j)th rows and columns, and that N_{i+j-2} is the common (i+j-2)th determinantal divisor. N' will not have the factor $g^{t_{ij}}$ since the rows and columns containing a'_{ii} and a'_{jj} were

deleted in forming M' and N' but will contain as factors all the remaining powers of g that occur in M_{i+j-1}.

Now $c'_{ij}N_{i+j-2}$ is a minor of order i+j-1 of N and as a consequence is a multiple of

$$\Pi_{k=1}^{i+j-1} g^{t_k} \tag{4.14}$$

which is a factor of M_{i+j-1}. Therefore, c'_{ij} must be a multiple of $g^{t_{ij}}$, the highest power of g which is common to a'_{ii} and a'_{jj}. Since g is any factor irreducible in $\mathcal{J}(x)$ and a factor of $a'_{\alpha\alpha}$, and since t_{ij} is the greatest power of g common to a'_{ii} and a'_{jj}, it follows that the greatest common factor, g_{ij} ($g_{ii} = a'_{ii}$) of a'_{ii} and a'_{jj} is a divisor of c'_{ij}. Consequently Eqs (4.8) are satisfied by elements u_{ij} in $\mathcal{J}(x)$ where $1 \leq i \leq \alpha$, $1 \leq j \leq \alpha$.

Case 2. The block of elements c'_{ij}, $1 \leq i \leq \alpha$, $\alpha < j \leq n$ is now examined. Here

$$c'_{ij}\Pi_{h=1}^{i-1}a'_{hh}\Pi_{h=i+1}^{\alpha}a'_{hh}\Pi_{k=1}^{\alpha}a'_{kk} \tag{4.15}$$

is a minor of order 2α of N; it must also be a multiple of

$$\Pi_{h=1}^{\alpha}a'_{hh}\Pi_{k=1}^{\alpha}a'_{kk} \tag{4.16}$$

which is the (2α)th determinantal divisor of M (of N). That is, c'_{ij} must be a multiple of a'_{ii} and Eqs (4.8) are satisfied for Case 2 by

$$u_{ij} = \frac{c'_{ij}}{a'_{ii}} \tag{4.17}$$

since a'_{jj} may be regarded as identically equal to zero where $\alpha < j \leq n$.

Case 3. Similarly Eqs (4.8) are valid for the block of elements c'_{ij}, $\alpha < i \leq n$, $1 \leq j \leq \alpha$. A similar type proof as in Case 2 holds.

Case 4. Finally,

$$c'_{ij} \equiv 0, \qquad \alpha < i \leq n, \ \alpha < j \leq n \tag{4.18}$$

otherwise N would have the nonzero minor

$$c'_{ij} \Pi_{h=1}^{\alpha} a'_{hh} \Pi_{k=1}^{\alpha} a'_{kk} \tag{4.19}$$

of order $2\alpha+1$, which is impossible. Hence, c'_{ij} is identically zero and any elements u_{ij} in $\mathcal{J}(x)$ will satisfy Eqs (4.8) since both a'_{ii} and a'_{jj} may be regarded as identically zero in case $\alpha < i \leq n$ and $\alpha < j \leq n$. Equations (4.8) are, therefore, valid for $i, j = 1, 2, \ldots, n$.

Therefore, $U = \{u_{ij}\}$ is a n by n matrix with elements in $\mathcal{J}(x)$ which satisfy Eq (4.7). Now $A' = PAQ$ and $C' = PCQ$ and since P and Q are non-singular, Eq (4.7) reduces to

$$\begin{aligned} [PAQ]U &= [PCQ] \\ A[QUQ^{-1}] &= C \\ X &= QUQ^{-1} \end{aligned} \tag{4.20}$$

and

with elements in $\mathcal{J}(x)$ such that Eq (4.1) is satisfied and the theorem is proved.

Example 4.1. The following example illustrates Theorem 4.2 for the matrix equation

$$AX = C \tag{4.21}$$

where

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 1+(\lambda-1)(\lambda+1) & 0 & 0 & 0 \\ 1 & -1-(\lambda-1)(\lambda+1) & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^4 & 0 & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)(\lambda+1)^2(\lambda+2)^5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.22}$$

and

$$C = \begin{bmatrix} 1 & -\lambda^2 & 0 & 0 & 0 \\ -1 & \lambda^2-\lambda^2(\lambda-1)(\lambda+1) & 0 & 0 & 0 \\ 1 & -\lambda^2-\lambda^2(\lambda-1)(\lambda+1) & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^5 & 0 & 0 \\ 0 & 0 & 0 & -\lambda^3(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.23}$$

By using

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.24}$$

$$A' = PAQ = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & (\lambda-1)(\lambda+1) & 0 & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^4 & 0 & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.25}$$

and

$$C' = PCQ = \begin{bmatrix} 1 & 1-\lambda^2 & 0 & 0 & 0 \\ 0 & \lambda^2(\lambda-1)(\lambda+1) & 0 & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^5 & 0 & 0 \\ 0 & 0 & 0 & -\lambda^3(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.26}$$

Taking A' and C' to form M

$$M = \left[\begin{array}{ccccc|} 1 & 0 & 0 & 0 & 0 \\ 0 & (\lambda-1)(\lambda+1) & 0 & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^4 & 0 & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$
\left[\begin{array}{ccccc|ccccc}
1 & 1-\lambda^2 & 0 & 0 & 0 & & & & & \\
0 & \lambda^2(\lambda-1)(\lambda+1) & 0 & 0 & 0 & & & & & \\
0 & 0 & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^5 & 0 & 0 & & & & & \\
0 & 0 & 0 & -\lambda^3(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 & & & & & \\
0 & 0 & 0 & 0 & 0 & & & & & \\
\hline
1 & 0 & 0 & 0 & 0 & & & & & \\
0 & (\lambda-1)(\lambda+1) & 0 & 0 & 0 & & & & & \\
0 & 0 & \lambda(\lambda-1)^2(\lambda+1)(\lambda+2)^4 & 0 & 0 & & & & & \\
0 & 0 & 0 & \lambda^2(\lambda-1)^3(\lambda+1)^2(\lambda+2)^5 & 0 & & & & & \\
0 & 0 & 0 & 0 & 0 & & & & &
\end{array}\right] \tag{4.27}
$$

Using

$$
a'_{ii}\, u_{ij} = c'_{ij}, \qquad i, j=1, 2, \ldots, 5 \tag{4.8}
$$

$$
U = \begin{bmatrix}
1 & 1-\lambda^2 & 0 & 0 & 0 \\
0 & \lambda^2 & 0 & 0 & 0 \\
0 & 0 & \lambda+2 & 0 & 0 \\
0 & 0 & 0 & -\lambda & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{4.28}
$$

and

$$
X = QUQ^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & \lambda^2 & 0 & 0 & 0 \\
0 & 0 & \lambda+2 & 0 & 0 \\
0 & 0 & 0 & -\lambda & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{4.29}
$$

which is the solution of Eq (4.21).

V. Solutions of Riccati Matrix Equations

Of all nonlinear differential equations, the Riccati differential equation

$$y'(z) = f_0(z) + f_1(z)y(z) + f_2(z)[y(z)]^2 \tag{5.1}$$

can claim to be the most important. It has a close connection with the second order linear case, moveable singularities which appear as simple poles, and a general solution that is a fractional linear function of the constant of integration (Ref 26:273). Count Jacopo Francesco Riccati lived in Italy from 1676 to 1754. He studied Newton's theories of gravitation and published works on philosphy, differential equations, and mensuration physics. His work with the Riccati equation culminated in 1724 when he published solutions to several special cases. This work was done before the introduction of matrices, however, and the term matrix Riccati equation remains somewhat of a misnomer.

One way the matrix Riccati equation originates is the following (Ref 23:237-239):

A time varying system represented by

$$\dot{\bar{x}} = A\bar{x} + B\bar{u} \tag{5.2}$$

is associated with a linear regulator problem with quadratic performance criterion

$$1/2\bar{x}^T M\bar{x} + 1/2\int_0^{t_1}[\bar{x}^T Q\bar{x} + \bar{u}^T R\bar{u}]dt \tag{5.3}$$

The Hamiltonian is

$$H = 1/2\bar{x}^T Q\bar{x} + 1/2\bar{u}^T R\bar{u} + \bar{p}^T(A\bar{x} + B\bar{u}) \tag{5.4}$$

and for optimality $\frac{\partial H}{\partial \bar{u}} = 0$ so that

$$\bar{u}^* = -R^{-1}B^T\bar{p}^* \tag{5.5}$$

The adjoint equations are

$$\dot{\bar{p}}^* = -Q\bar{x}^* - A^T p^* \tag{5.6}$$

and by substitution

$$\dot{\bar{x}}^* = A\bar{x}^* - BR^{-1}B^T\bar{p}^* \tag{5.7}$$

The solution of Eq (5.6) can be written in terms of a transition matrix P:

$$\bar{p}^* = P\bar{x}^* \tag{5.8}$$

To determine P, Eq (5.8) is differentiated to get

$$\dot{P}\bar{x}^* + P\dot{\bar{x}}^* - \dot{\bar{p}}^* = 0 \tag{5.9}$$

and by substituting for $\dot{\bar{x}}^*$, $\dot{\bar{p}}^*$, and $\bar{p}^*$

$$(\dot{P} + PA - PBR^{-1}B^T P + Q + A^T P)\bar{x}^* = 0 \tag{5.10}$$

Since Eq (5.10) must be valid for all time $0 \leq t \leq t_1$, P must satisfy

$$\dot{P} = PBR^{-1}B^T P - A^T P - PA - Q \tag{5.11}$$

This is the form called the matrix Riccati equation and the steady state ($\dot{P} = 0$) form is referred to as the algebraic matrix Riccati equation.

The algebraic matrix Riccati equation in the form

$$XDX + AX + XB + C = 0 \tag{5.12}$$

where A, B, C, and D are n by n matrices having elements belonging to the field of complex numbers is considered in this chapter. Results of W. E. Roth (Ref 27) are extended. The coefficients of all polynomials which arise in this chapter belong to the field of complex numbers. Therefore, the similarity of matrices and the reduceability of polynomials remain valid under the rational operations of the field.

5.1. Necessary Conditions for the Existence of Solutions of Eq (5.12)

If the 3n by 3n matrix $\tilde{R}$ is given by

$$\tilde{R} = \begin{bmatrix} -B & 0 & D \\ 0 & I & 0 \\ -C & 0 & A \end{bmatrix} \tag{5.13}$$

and if X is any solution of Eq (5.12) with elements belonging to the field of complex numbers, then the following eight theorems result:

Theorem 5.1. If X is any solution of Eq (5.12), then the following matrices

$$\tilde{R} = \begin{bmatrix} -B & 0 & D \\ 0 & I & 0 \\ -C & 0 & A \end{bmatrix},$$

$$\begin{bmatrix} XD+A & 0 & (-XB-C)+(XD+A)(-X) \\ 0 & I & 0 \\ D & 0 & -DX-B \end{bmatrix} = \begin{bmatrix} XD+A & 0 & 0 \\ 0 & I & 0 \\ D & 0 & -DX-B \end{bmatrix} = \tilde{\tilde{R}} \tag{5.14}$$

are similar, and there exists at least one triple of polynomials $f_\alpha(\lambda)$ of degree $\alpha \leq n$, $g_\beta(\lambda)$ of degree $\beta \leq n$, and $h_\gamma(\lambda)$ of degree $\gamma \leq n$ with coefficients in the field of complex numbers, such that $f_\alpha(\tilde{R}) \cdot g_\beta(\tilde{R}) \cdot h_\gamma(\tilde{R}) = 0$, where $f_\alpha(\lambda) \cdot g_\beta(\lambda) \cdot h_\gamma(\lambda)$ is not necessarily the minimum polynomial

satisfied by $\tilde{R}$ but is a divisor of $|\tilde{R} - \lambda I|$ such that $f_\alpha(XD + A) = 0$, $g_\beta(I) = 0$, and $h_\gamma(-B - DX) = 0$.

Proof. If X is a solution of Eq (5.12), then

$$\begin{bmatrix} X & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{bmatrix} \begin{bmatrix} -B & 0 & D \\ 0 & I & 0 \\ -C & 0 & A \end{bmatrix} \begin{bmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & -X \end{bmatrix} = \begin{bmatrix} XD+A & 0 & (-XB-C)+(XD+A)(-X) \\ 0 & I & 0 \\ D & 0 & -DX-B \end{bmatrix} = \tilde{\tilde{R}} \tag{5.15}$$

Therefore, $\tilde{R}$ and $\tilde{\tilde{R}}$ are similar. Now $|\tilde{R} - \lambda I| = |XD + A - \lambda I| \cdot |I - \lambda I| \cdot |-DX - B - \lambda I|$. If $f_{\tilde{R}}(\lambda) = |\tilde{R} - \lambda I| = f_\alpha(\lambda) \cdot g_\beta(\lambda) \cdot h_\gamma(\lambda)$ then $0 = f_\alpha(XD + A) \cdot g_\beta(I) \cdot h_\gamma(-DX - B)$ and at least one polynomial must vanish.

Polynomials $f_\alpha(\lambda)$, $g_\beta(\lambda)$, $h_\gamma(\lambda)$ of degree $\alpha \leq n$, $\beta \leq n$, $\gamma \leq n$ respectively such that $f_\alpha(\lambda) \cdot g_\beta(\lambda) \cdot h_\gamma(\lambda)$ is a divisor of $|\tilde{R} - \lambda I|$ and a multiple of the minimum polynomial satisfied by $\tilde{R}$ are called an admissible class of polynomials and are denoted by $\mathcal{A}(R)$.

Theorem 5.2 (Ref 28). If $f_\alpha(\lambda) \in \mathcal{A}(\tilde{R})$ such that $f_\alpha(X_1 D + A) = 0$ where X_1 is any solution of Eq (5.12), and if

$$f_\alpha(\tilde{R}) = \begin{bmatrix} U & \bar{0} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \tag{5.16}$$

where $\tilde{R}$ is given in Eq (5.14) and where U, V, M, N, $\bar{0}$, P, Q, R, and S are polynomials in the matrices A, B, C, and D, then the following equations:

$$X_1 U + V = 0, \qquad X_1 M + N = 0, \qquad X_1 \bar{0} + S = 0 \tag{5.17}$$

have a common solution X_1 which also satisfies Eq (5.12). In addition, the following holds:

$$(X_1 \quad 0 \quad I) f_\alpha(\tilde{R}) = (0 \quad 0 \quad 0) \tag{5.18}$$

Proof. If X_1 is a solution of Eq (5.12) and if $f_\alpha(\tilde{R})$ is given by Eq (5.16), then using Eq (5.15),

$$\begin{bmatrix} X_1 & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{bmatrix} f_\alpha(\tilde{R}) \begin{bmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & -X_1 \end{bmatrix} = \begin{bmatrix} f_\alpha(X_1D+A) & 0 & 0 \\ * & f_\alpha(I) & 0 \\ ** & *** & f_\alpha(-DX_1-B) \end{bmatrix}$$

$$= \begin{bmatrix} X_1M+N & X_1\bar{O}+S & (X_1U+V)+(X_1M+N)(-X_1) \\ R & Q & P-RX_1 \\ M & \bar{O} & U-MX_1 \end{bmatrix} \tag{5.19}$$

where *, **, *** denote polynomials in A, B, C, and D. Since $f_\alpha(X_1D + A) = 0$ then $X_1M + N = 0$, $X_1\bar{O} + S = 0$, and $X_1U + V = 0$ by matching 0 components of Eq (5.19). Equation (5.18) follows immediately.

Theorem 5.3. If $h_\gamma(-DX_2 - B) = 0$ such that $h_\gamma(\lambda) \in \mathcal{A}(\tilde{R})$ where X_2 is any solution of Eq (5.12), and if

$$h_\gamma(\tilde{R}) = \begin{bmatrix} \bar{U} & \bar{\bar{O}} & \bar{M} \\ \bar{P} & \bar{Q} & \bar{R} \\ \bar{V} & \bar{S} & \bar{N} \end{bmatrix} \tag{5.20}$$

where $\tilde{R}$ is given in Eq (5.14) and where $\bar{U}$, $\bar{V}$, $\bar{M}$, $\bar{N}$, $\bar{\bar{O}}$, $\bar{P}$, $\bar{Q}$, $\bar{R}$, and $\bar{S}$ are polynomials in the matrices A, B, C, and D, then the following equations:

$$\bar{U} - \bar{M}X_2 = 0, \qquad \bar{V} - \bar{N}X_2 = 0, \qquad \bar{P} - \bar{R}X_2 = 0 \tag{5.21}$$

have a common solution X_2 which also satisfies Eq (5.12), In addition,

the following holds:

$$h_\gamma(\tilde{R}) \begin{bmatrix} I \\ 0 \\ -X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.22}$$

Proof. If X_2 is a solution of Eq (5.12) and if $h_\gamma(\tilde{R})$ is given by Eq (5.20), then using Eq (5.15),

$$\begin{bmatrix} X_2 & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{bmatrix} h_\gamma(\tilde{R}) \begin{bmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & -X_2 \end{bmatrix} = \begin{bmatrix} h_\gamma(X_2D+A) & 0 & 0 \\ * & h_\gamma(I) & 0 \\ ** & *** & h_\gamma(-DX_2-B) \end{bmatrix}$$
$$= \begin{bmatrix} X_2\bar{M}+\bar{N} & X_2\bar{\bar{O}}+\bar{S} & (X_2\bar{U}+\bar{V})+(X_2\bar{M}+\bar{N})(-X_2) \\ \bar{R} & \bar{Q} & \bar{P}-\bar{R}X_2 \\ \bar{M} & \bar{\bar{O}} & \bar{U}-\bar{M}X_2 \end{bmatrix} \tag{5.23}$$

where *, **, *** denote polynomials in A, B, C, and D. Since $h_\gamma(-DX_2-B) = 0$ then $\bar{U} - \bar{M}X_2 = 0$, $\bar{P} - \bar{R}X_2 = 0$ by matching 0 components of Eq (5.23). Now $\bar{U} = \bar{M}X_2$ and $(X_2\bar{U} + \bar{V}) + (X_2\bar{M} + \bar{N})(-X_2) = 0$ implies $X_2\bar{U} + \bar{V} - X_2\bar{M}X_2 - \bar{N}X_2 = 0$ implies $X_2\bar{U} + \bar{V} - X_2\bar{U} - \bar{N}X_2 = 0$ implies $\bar{V} - \bar{N}X_2 = 0$. Equation (5.22) follows immediately.

Theorem 5.4. If $g_\beta(\lambda) \in \mathcal{A}(\tilde{R})$ such that $g_\beta(I) = 0$ where X_3 is any solution of Eq (5.12), and if

$$g_\beta(\tilde{R}) = \begin{bmatrix} \bar{\bar{U}} & \bar{\bar{\bar{O}}} & \bar{\bar{M}} \\ \bar{\bar{P}} & \bar{\bar{Q}} & \bar{\bar{R}} \\ \bar{\bar{V}} & \bar{\bar{S}} & \bar{\bar{N}} \end{bmatrix} \tag{5.24}$$

where $\tilde{R}$ is given in Eq (5.14) and where $\bar{\bar{U}}$, $\bar{\bar{V}}$, $\bar{\bar{M}}$, $\bar{\bar{N}}$, $\bar{\bar{\bar{O}}}$, $\bar{\bar{P}}$, $\bar{\bar{Q}}$, $\bar{\bar{R}}$, and $\bar{\bar{S}}$ are polynomials in the matrices A, B, C, and D, then the following equations:

$$\overline{\overline{Q}} = 0, \qquad X_3\overline{\overline{O}} + S = 0, \qquad \overline{\overline{P}} - \overline{\overline{R}}X_3 = 0 \tag{5.25}$$

$$(X_3\overline{\overline{U}}+\overline{\overline{V}}) + (X_3\overline{\overline{M}}+\overline{\overline{N}})(-X_3) = 0 \tag{5.26}$$

have a common solution X_3 which also satisfies Eq (5.12).

Proof. If X_3 is a solution of Eq (5.12), and if $g_\beta(\tilde{R})$ is given by Eq (5.24), then using Eq (5.15),

$$\begin{bmatrix} X_3 & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{bmatrix} g_\beta(\tilde{R}) \begin{bmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & -X_3 \end{bmatrix} = \begin{bmatrix} g_\beta(X_3D+A) & 0 & 0 \\ * & g_\beta(I) & 0 \\ ** & *** & g_\beta(-DX_3-B) \end{bmatrix}$$

$$= \begin{bmatrix} X_3\overline{\overline{M}}+\overline{\overline{N}} & X_3\overline{\overline{O}}+\overline{\overline{S}} & (X_3\overline{\overline{U}}+\overline{\overline{V}})+(X_3\overline{\overline{M}}+\overline{\overline{N}})(-X_3) \\ \overline{\overline{R}} & \overline{\overline{Q}} & \overline{\overline{P}}-\overline{\overline{R}}X_3 \\ \overline{\overline{M}} & \overline{\overline{O}} & \overline{\overline{U}}-\overline{\overline{M}}X_3 \end{bmatrix} \tag{5.27}$$

where *, **, *** denote polynomials in A, B, C, and D. Since $g_\beta(I) = 0$ then $(X_3\overline{\overline{U}} + \overline{\overline{V}}) + (X_3\overline{\overline{M}} + \overline{\overline{N}})(-X_3) = 0$, $\overline{\overline{Q}} = 0$, $X_3\overline{\overline{O}} + \overline{\overline{S}} = 0$, and $\overline{\overline{P}} - \overline{\overline{R}}X_3 = 0$ by matching 0 components of Eq (5.27).

Theorem 5.5. If $f_\delta(\lambda) \in \mathcal{A}(\tilde{R})$, and if

$$(X \quad 0 \quad I) f_\delta(\tilde{R}) = (0 \quad 0 \quad 0) \text{ and } f_\delta(\tilde{R}) \begin{bmatrix} I \\ 0 \\ -X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.28}$$

then

$$f_\delta(\tilde{R}) = \begin{bmatrix} \hat{U} & \hat{O} & \hat{M} \\ \hat{P} & \hat{Q} & \hat{R} \\ \hat{V} & \hat{S} & \hat{N} \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & 0 \\ \hat{R} & \hat{Q} & 0 \\ \hat{M} & \hat{O} & 0 \end{bmatrix} = \overline{R} \tag{5.29}$$

are similar.

Proof. If X satisfies Eqs (5.28), then

$$\begin{bmatrix} X & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{U} & \hat{O} & \hat{M} \\ \hat{P} & \hat{Q} & \hat{R} \\ \hat{V} & \hat{S} & \hat{N} \end{bmatrix} \begin{bmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & -X \end{bmatrix} = \begin{bmatrix} \hat{XM}+\hat{N} & \hat{XO}+\hat{S} & (\hat{XU}+\hat{V})+(\hat{XM}+\hat{N})(-X) \\ \hat{R} & \hat{Q} & \hat{P}-\hat{RX} \\ \hat{M} & \hat{O} & \hat{U}-\hat{MX} \end{bmatrix} \tag{5.30}$$

If X_1 is a solution of the first equation of Eq (5.28) and X_2 is a solution of the second equation of Eq (5.28), then the matrices in Eq (5.29) are similar.

The following identities implied by $\tilde{R}f_\alpha(\tilde{R}) = f_\alpha(\tilde{R})\tilde{R}$ are developed now for later use in this chapter:

$$\begin{bmatrix} -B & 0 & D \\ 0 & I & 0 \\ -C & 0 & A \end{bmatrix} \begin{bmatrix} U & \bar{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} = \begin{bmatrix} U & \bar{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \begin{bmatrix} -B & 0 & D \\ 0 & I & 0 \\ -C & 0 & A \end{bmatrix} \tag{5.31}$$

$$\begin{bmatrix} -BU+DV & -B\bar{O}+DS & -BM+DN \\ P & Q & R \\ -CU+AV & -C\bar{O}+AS & -CM+AN \end{bmatrix} = \begin{bmatrix} -UB-MC & \bar{O} & UD+MA \\ -PB-RC & Q & PD+RA \\ -VB-NC & S & VD+NA \end{bmatrix} \tag{5.32}$$

$$\begin{aligned} -BU + DV &= -UB - MC, & -B\bar{O} + DS &= \bar{O}, & -BM + DN &= UD + MA \\ P &= -PB - RC, & Q &= Q, & R &= PD + RA \\ -CU + AV &= -VB - NC, & -C\bar{O} + AS &= S, & -CM + AN &= VD + NA \end{aligned} \tag{5.33}$$

Next, solutions of the following pair of equations are related

$$(X \quad 0 \quad I)f_\alpha(\tilde{R}) = (0 \quad 0 \quad 0) \tag{5.34}$$

and

$$f_\alpha(\tilde{R}) \begin{bmatrix} I \\ 0 \\ -X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.35}$$

$$(X \quad 0 \quad I)\begin{bmatrix} U & \bar{0} & M \\ P & Q & R \\ V & S & N \end{bmatrix} = (0 \quad 0 \quad 0), \qquad \begin{bmatrix} U & \bar{0} & M \\ P & Q & R \\ V & S & N \end{bmatrix}\begin{bmatrix} I \\ 0 \\ -X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.36}$$

$$\begin{aligned} XU + V &= 0 & \qquad U - MX &= 0 \\ X\bar{0} + S &= 0 & \qquad P - RX &= 0 \\ XM + N &= 0 & \qquad V - NX &= 0 \end{aligned} \tag{5.37}$$

Theorem 5.6. If X_1 is a solution of Eq (5.34) and $f_\alpha(\lambda) \in \mathcal{A}(R)$ such that

$$f_\alpha(\tilde{R}) = \begin{bmatrix} U & \bar{0} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \tag{5.38}$$

and M is nonsingular, then

$$X_2D + A = M^{-1}(-DX_1 - B)M \tag{5.39}$$

It will be shown that a unique solution of Eq (5.35) exists and that

$$U - MX_2 = 0, \qquad V - NX_2 = 0 \tag{5.40}$$

Also, if X_1 is a solution of Eq (5.34), it will be shown that

$$X_1U + V = 0, \qquad X_1M + N = 0 \tag{5.41}$$

Proof. Using $X_1M + N = 0$, $X_1U + V = 0$, $U - MX_2 = 0$, and $V - NX_2 = 0$ it follows that $X_2 = M^{-1}U$ and $X_2D = M^{-1}UD$. Next, from the identities in Eq (5.33)

$$
\begin{aligned}
X_2D &= M^{-1}(DN - MA - BM) \\
&= M^{-1}\,DN - A - M^{-1}BM \\
&= M^{-1}(-DX_1M) - M^{-1}BM - A \\
X_2D + A &= M^{-1}(-DX_1 - B)M
\end{aligned}
\tag{5.42}
$$

Equation (5.39) follows.

Theorem 5.7. If X_1 is a solution of Eq (5.34) and $f_\alpha(\lambda) \in \mathcal{A}(\tilde{R})$ such that

$$
f_\alpha(\tilde{R}) = \begin{bmatrix} U & \bar{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \tag{5.43}
$$

and U is nonsingular then

$$
DX_1 + B = -M(C + X_2B)U^{-1} \tag{5.44}
$$

Proof. If U^{-1} exists, then $X_1 = -VU^{-1}$ and $DX_1 = -DVU^{-1}$, then

$$
\begin{aligned}
DX_1 &= -(BU - MC - UB)U^{-1} \\
&= -(B - MCU^{-1} - UBU^{-1}) \\
&= -(B - MCU^{-1} - MX_2BU^{-1}) \\
DX_1 + B &= M(C + X_2B)U^{-1}
\end{aligned}
\tag{5.45}
$$

Theorem 5.8. If X_1 is a solution of Eq (5.34) and $f_\alpha(\lambda) \in \mathcal{A}(R)$ such that

$$
f_\alpha(\tilde{R}) = \begin{bmatrix} U & \bar{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \tag{5.46}
$$

and N is nonsingular then

$$X_2D + A = -N^{-1}(AX_1 + C)M \tag{5.47}$$

Proof. If N^{-1} exists then $V = NX_2$ and

$$\begin{aligned} X_2D &= N^{-1}VD \\ &= N^{-1}(AN - NA - CM) \\ &= N^{-1}AN - A - N^{-1}CM \\ &= -N^{-1}AX_1M - A - N^{-1}CM \\ X_2D + A &= -N^{-1}(AX_1 + C)M \end{aligned} \tag{5.48}$$

5.2. Sufficient Conditions for the Existence of Solutions of Eq (5.12)

In this section f is a complex-valued function of a complex variable which is defined on the spectrum of $\tilde{R}$ such that $\tilde{R}f(\tilde{R}) = f(\tilde{R})\tilde{R}$ (Ref 27), where

$$f(\tilde{R}) = \begin{bmatrix} U & \overline{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix}, \qquad \tilde{R} = \begin{bmatrix} -B & O & D \\ O & I & O \\ -C & O & A \end{bmatrix} \tag{5.49}$$

Theorem 5.9. If M^{-1} or N^{-1} exists and X is a solution of

$$f(\tilde{R}) \begin{bmatrix} I \\ O \\ -X \end{bmatrix} = \begin{bmatrix} O \\ O \\ O \end{bmatrix} \tag{5.50}$$

where R^{-1} exists, then the pair of equations

$$\begin{aligned} A + XD &= I \\ AX + XB + C + XDX &= O \end{aligned} \tag{5.51}$$

has X as a common solution given by

$$X = (I - A)D^{-} + Z - ZDD^{-} \tag{5.52}$$

for any arbitrary Z (Ref 20).

Proof.

$$f(\tilde{R}) \begin{bmatrix} I \\ O \\ -X \end{bmatrix} = \begin{bmatrix} U & \overline{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} \begin{bmatrix} I \\ O \\ -X \end{bmatrix} = \begin{bmatrix} O \\ O \\ O \end{bmatrix} \tag{5.53}$$

$$U - MX = O, \qquad P - RX = O, \qquad V - NX = O \tag{5.54}$$

Using R = PD + RA from the matrix identities in Eq (5.33) and P = RX from Eq (5.54)

$$\begin{aligned} -R + PD + RA &= O \\ -R + RXD + RA &= O \\ R[-I + XD + A] &= O \end{aligned} \tag{5.55}$$

Since R^{-1} exists

$$\begin{aligned} O &= -I + XD + A \\ I &= A + XD \\ X &= (I - A)D^{-} + Z - ZDD^{-} \end{aligned} \tag{5.56}$$

and using V - NX = O from Eq (5.54) plus the matrix identities in Eq (5.33)

$$\begin{aligned} O &= D(V - NX) \\ &= DV - DNX \\ &= (-UB + BU - MC) - (UD + MA + BM)X \end{aligned} \tag{5.57}$$

$$= (-MXB + BMX - MC) - (MXD + MA + BM)X$$

$$= -MXB + BMX - MC - MXDX - MAX - BMX$$

$$= -M(XB + AX + C + XDX)$$

Since M^{-1} exists this implies that X is a solution of Eq (5.12). Similarly,

$$\begin{aligned} 0 &= A(V - NX) \\ &= AV - ANX \\ &= (-VB - NC + CU) - (VD + NA + CM)X \\ &= -NXB - NC + CMX - NXDX - NAX - CMX \\ &= -N(XB + AX + C + XDX) \end{aligned} \tag{5.58}$$

Since N^{-1} exists this implies that X is a solution of Eq (5.12).

Theorem 5.10. If M^{-1} or U^{-1} exist and X is a solution of

$$(X \quad 0 \quad I)f(\tilde{R}) = (0 \quad 0 \quad 0) \tag{5.59}$$

where $\bar{O}^{-1}$ exists, then the pair of equations

$$\begin{aligned} -I &= B + DX \\ 0 &= AX + XB + C + XDX \end{aligned} \tag{5.60}$$

has X as a common solution given by

$$X = \bar{D}(-I - B) + W - \bar{D}D \tag{5.61}$$

for any arbitrary W (Ref 20).

Proof.

$$(X \quad 0 \quad I)f(\tilde{R}) = (X \quad 0 \quad I)\begin{bmatrix} U & \bar{O} & M \\ P & Q & R \\ V & S & N \end{bmatrix} = (0 \quad 0 \quad 0) \tag{5.62}$$

$$(XU + V \quad X\bar{O} + S \quad XM + N) = (O \quad O \quad O) \tag{5.63}$$

Using $-B\bar{O} + DS = \bar{O}$ from the matrix identities in Eq (5.33) and $X\bar{O} + S = O$ from Eq (5.63)

$$\begin{aligned} \bar{O} + B\bar{O} - DS &= O \\ \bar{O} + B\bar{O} + DX\bar{O} &= O \\ [I + B + DX]\bar{O} &= O \end{aligned} \tag{5.64}$$

Since $\bar{O}^{-1}$ exists

$$\begin{aligned} O &= I + B + DX \\ -I &= B + DX \\ X &= D^{-}(-I - B) + W - D^{-}D \end{aligned} \tag{5.65}$$

for any arbitrary W, and using $XM + N = O$ from Eq (5.63) plus the matrix identities in Eq (5.33)

$$\begin{aligned} O &= (XM + N)A \\ &= XMA + NA \\ &= -VD + XMA + NA + VD \\ &= XUD + XMA + NA + VD \\ &= X(UD + MA) + NA + VD \\ &= X(-BM + DN) + NA + VD \\ &= -XBM + XDN - AXM + AXM + NA + VD \\ &= -XBM - XDXM - AXM - AN + NA + VD \\ &= -XBM - XDXM - AXM - CM \\ &= -(XB + XDX + AX + C)M \end{aligned} \tag{5.66}$$

Since M^{-1} exists this implies that X is a solution of Eq (5.12). Similarly,

$$
\begin{aligned}
0 &= -(XM + N)C \\
&= -XMC - NC \\
&= -XUB - XMC + XUB - NC \\
&= -X(UB + MC) + XUB - NC \\
&= X(DV - BU) + XUB - NC \\
&= XDV - XBU + XUB - NC \qquad (5.67) \\
&= -XDXU - XBU + XUB - NC \\
&= -XDXU - XBU - AXU + AXU + XUB - NC \\
&= -XDXU - XBU - AXU - AV - VB - NC \\
&= -XDXU - XBU - AXU - CU \\
&= -(XDX + XB + AX + C)U
\end{aligned}
$$

Since U^{-1} exists this implies that X is a solution of Eq (5.12).

5.3. Uniqueness Theorems

Theorem 5.11. If the following conditions hold

(i) $DA = BD$

(ii) $(-A \pm I) = (-A \pm I)\bar{D}D$

(iii) $(-B \pm I) = D\bar{D}(-B \pm I)$

(iv) Sufficient conditions that $XDX - 2/3X + 1/3C = 0$ have a solution

then

$$0 = XDX + AX + XB + C \qquad (5.68)$$

$$A = -XD \pm I \qquad (5.69)$$

$$B = -DX \pm I \qquad (5.70)$$

have a unique solution X. Such a solution also satisfies

$$XDX - 2/3X + 1/3C = 0 \qquad (5.71)$$

Proof. Conditions (i), (ii), and (iii) from Mitra (Ref 20) insure the Eqs (5.69) and (5.70). Condition (iv) is satisfied by Theorems 5.9 and 5.10. In addition, if X_1 and X_2 are solutions of Eq (5.68) where $X_1 \neq X_2$, then

$$\begin{aligned} AX_1 + X_1B + C + X_1DX_1 &= 0 \\ AX_2 + X_2B + C + X_2DX_2 &= 0 \end{aligned} \tag{5.72}$$

subtracting, $A(X_1 - X_2) + (X_1 - X_2)B + X_1DX_1 - X_2DX_2 = 0$

$$\begin{aligned} A(X_1 - X_2) + (X_1 - X_2)B + X_1DX_1 - X_2DX_1 + X_2DX_1 - X_2DX_2 &= 0 \\ A(X_1 - X_2) + (X_1 - X_2)B + (X_1 - X_2)DX_1 + X_2D(X_1 - X_2) &= 0 \\ (A + X_2D)(X_1 - X_2) + (X_1 - X_2)(B + DX_1) &= 0 \end{aligned} \tag{5.73}$$

Since

$$\begin{aligned} A + XD &= \pm I \\ B + DX &= \pm I \end{aligned} \tag{5.74}$$

then $\pm I(X_1 - X_2) + (X_1 - X_2)(\pm I) = 0$

$$2X_1 = 2X_2 \tag{5.75}$$

and $X_1 = X_2$

which contradicts the hypothesis and, therefore, the solution is unique.

Theorem 5.12. If M^{-1} or N^{-1} exists and X is a solution of

$$f(\tilde{R}) \begin{bmatrix} I \\ 0 \\ -X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.76}$$

where R^{-1} exists, then the pair of equations

$$A^T + XD = I \tag{5.77}$$

$$A^T X + XA + C + XDX = 0 \tag{5.78}$$

has X as a common solution. In optimal control theory where D is positive definite, $D = D^T$, and $X = X^T$, such a solution is unique. Also, the optimal control theory notation corresponds to $A^T = A$ and $A = B$.

Proof. By Theorem 5.11, Eqs (5.77) and (5.78) have X as a common solution. If X_1 and X_2 are solutions to the Riccati equation where $X_1 \neq X_2$,

$$\begin{aligned} A^T X_1 + X_1 A + X_1 D X_1 + C &= 0 \\ A^T X_2 + X_2 A + X_2 D X_2 + C &= 0 \end{aligned} \tag{5.79}$$

subtracting, $A^T(X_1 - X_2) + (X_1 - X_2)A + X_1 DX_1 - X_2 DX_2 = 0$

$$\begin{aligned} A^T(X_1 - X_2) + (X_1 - X_2)A + X_1 DX_1 - X_1 DX_2 + X_1 DX_2 - X_2 DX_2 &= 0 \\ A^T(X_1 - X_2) + (X_1 - X_2)A + X_1 D(X_1 - X_2) + (X_1 - X_2)DX_2 &= 0 \\ (A^T + X_1 D)(X_1 - X_2) + (X_1 - X_2)(A + DX_2) &= 0 \end{aligned} \tag{5.80}$$

from Eq (5.77), $A^T + XD = I$ and since $D = D^T$ and $X = X^T$,

$$\begin{aligned} (A^T + XD)^T &= I^T \\ A + D^T X^T &= I \\ A + DX &= I \end{aligned} \tag{5.81}$$

which implies $I(X_1 - X_2) + (X_1 - X_2)I = 0$

$$\begin{aligned} X_1 - X_2 + X_1 - X_2 &= 0 \\ 2X_1 &= 2X_2 \end{aligned} \tag{5.82}$$

and $$X_1 = X_2$$

which contradicts the hypothesis and, therefore, only one unique solution exists.

Theorem 5.13. If M^{-1} or U^{-1} exists and X is solution of

$$(I \quad 0 \quad X)f(\tilde{R}) = (0 \quad 0 \quad 0) \tag{5.83}$$

where $\bar{O}^{-1}$ exists, then the pair of equations

$$A + DX = -I \tag{5.84}$$

$$A^T X + XA + C + XDX = 0 \tag{5.85}$$

has X as a common solution. In optimal control theory where D is positive definite, $D = D^T$, and $X = X^T$, such a solution is unique.

Proof. By Theorem 5.11, Eqs (5.84) and (5.85) have X as a common solution. If X_1 and X_2 are solutions to the Riccati equation where $X_1 \neq X_2$, and by using Eqs (5.79) and (5.80),

$$(A^T + X_1 D)(X_1 - X_2) + (X_1 - X_2)(A + DX_2) = 0 \tag{5.86}$$

from Eq (5.84) $A + DX = -I$ and since $D = D^T$ and $X = X^T$,

$$\begin{aligned} (A + DX)^T &= -I^T \\ A^T + X^T D^T &= -I \\ A^T + XD &= -I \end{aligned} \tag{5.87}$$

which implies

$$\begin{aligned} -I(X_1 - X_2) + (X_1 - X_2)(-I) &= 0 \\ X_2 - X_1 + X_2 - X_1 &= 0 \\ 2X_2 &= 2X_1 \end{aligned} \tag{5.88}$$

and

$$X_2 = X_1$$

which contradicts the hypothesis and, therefore, only one unique solution exists.

Example 5.1. An example from Martensson (Ref 29:23) is used to show the computational technique presented in the previous sections. Martensson's equation was written

$$A^T X + XA - XBQ_2^{-1}B^T X + Q_1 = 0 \tag{5.89}$$

where

$$A^T = \begin{bmatrix} -3 & -2 \\ 2 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} -3 & 2 \\ -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad B^T = [0 \quad 1],$$
$$Q_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q_2 = [1] \tag{5.90}$$

Rewriting in the form of Eq (5.12) where

$$D = -BQ_2^{-1}B^T = -\begin{bmatrix} 0 \\ 1 \end{bmatrix}[1][0 \quad 1] = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \tag{5.91}$$

$$X \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} X + \begin{bmatrix} -3 & -2 \\ 2 & 1 \end{bmatrix} X + X \begin{bmatrix} -3 & 2 \\ -2 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{5.92}$$

$$R = \begin{bmatrix} -B & D \\ -C & A \end{bmatrix} = \begin{bmatrix} -3 & -2 & 0 & 0 \\ 2 & -1 & 0 & -1 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & 2 & 1 \end{bmatrix} \tag{5.93}$$

$$|R-\lambda I| = (\lambda-1)^2(\lambda+1)^2 = 0 \tag{5.94}$$

TABLE I

Combinations of Characteristic Roots of R

	1	-1
1	$(\lambda-1)(\lambda-1)$	$(\lambda-1)(\lambda+1)$
-1	$(\lambda+1)(\lambda-1)$	$(\lambda+1)(\lambda+1)$

Table I shows three combinations of the characteristic roots of R to be checked:

(i) Case 1 $R^2 - 2R + I$

(ii) Case 2 $R^2 - I$

(iii) Case 3 $R^2 + 2R + I$

Case 1. $R^2 - 2R + I$

$$\begin{bmatrix} 5 & -4 & 0 & 2 \\ 4 & -3 & -2 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & -4 & -3 \end{bmatrix} + \begin{bmatrix} -6 & 4 & 0 & 0 \\ -4 & 2 & 0 & 2 \\ 0 & 0 & 6 & 4 \\ 0 & 0 & -4 & -2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 12 & 8 \\ 0 & 0 & -8 & -4 \end{bmatrix} = \begin{bmatrix} U & M \\ V & N \end{bmatrix} \qquad (5.95)$$

$$X = -NM^{-1} = -\begin{bmatrix} 12 & 8 \\ -8 & -4 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 0 \end{bmatrix} = \begin{bmatrix} -10 & 6 \\ 6 & -4 \end{bmatrix} \qquad (5.96)$$

Case 2. $R^2 - I$

$$\begin{bmatrix} 5 & -4 & 0 & 2 \\ 4 & -3 & -2 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & -4 & -3 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -4 & 0 & 2 \\ 4 & -4 & -2 & 0 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & -4 & -4 \end{bmatrix} = \begin{bmatrix} U & M \\ V & N \end{bmatrix} \tag{5.97}$$

$$X = M^{-1}U = \begin{bmatrix} 0 & -1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} 4 & -4 \\ 4 & -4 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \tag{5.98}$$

also,

$$X = -NM^{-1} = -\begin{bmatrix} 4 & 4 \\ -4 & -4 \end{bmatrix} \begin{bmatrix} 0 & -1/2 \\ 1/2 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \tag{5.99}$$

Case 3. The results are the same as Case 1.

Both solutions can be shown to satisfy Eqs (5.89) and (5.92).

Example 5.2. This example, taken from Denman and Beavers (Ref 30: 69), shows several solutions not found by other techniques.

$$A = \begin{bmatrix} 0 & 0 \\ 1 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \tag{5.100}$$

$$R = \begin{bmatrix} -B & D \\ -C & A \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & -1 \\ -4 & 0 & 0 & 0 \\ 0 & -1 & 1 & -2 \end{bmatrix} \tag{5.101}$$

$$|R-\lambda I| = (\lambda^2-4)(\lambda^2-1) = 0 \tag{5.102}$$

TABLE II

Combinations of Characteristic Roots of R

	1	-1	2	-2
1	$(\lambda-1)^2$	$(\lambda-1)(\lambda+1)$	$(\lambda-1)(\lambda-2)$	$(\lambda-1)(\lambda+2)$
-1	$(\lambda+1)(\lambda-1)$	$(\lambda+1)^2$	$(\lambda+1)(\lambda-2)$	$(\lambda+1)(\lambda+2)$
2	$(\lambda-2)(\lambda-1)$	$(\lambda-2)(\lambda+1)$	$(\lambda-2)^2$	$(\lambda-2)(\lambda+2)$
-2	$(\lambda+2)(\lambda-1)$	$(\lambda+2)(\lambda+1)$	$(\lambda+2)(\lambda-2)$	$(\lambda+2)^2$

Table II shows that ten possibilities exist in this example:

Case 1. $R^2 - 2R + I$ - No solutions

Case 2. $R^2 - I$

$$\begin{bmatrix} 0 & -2 & 0 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 4 & 0 & 0 \\ -4 & 0 & -2 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -2 & 0 & 1 \\ 0 & 4 & -1 & 0 \\ 0 & 4 & -1 & 0 \\ -4 & 0 & -2 & 4 \end{bmatrix} = \begin{bmatrix} U & M \\ V & N \end{bmatrix} \qquad (5.103)$$

$$X = M^{-1}U = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & -4 \\ -1 & -2 \end{bmatrix}$$

$$X = N^{-1}V = \begin{bmatrix} -1 & 0 \\ -1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -4 \\ -1 & -2 \end{bmatrix} \qquad (5.104)$$

and

$$X = -NM^{-1} = \begin{bmatrix} 1 & 0 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -4 & -2 \end{bmatrix}$$

$$X = -VU^{-1} = \begin{bmatrix} 0 & -4 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1/2 \\ 0 & 1/4 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -4 & -2 \end{bmatrix} \qquad (5.105)$$

Case 3. $R^2 - 3R + 2I$

$$X = \begin{bmatrix} 6 & 2 \\ 2 & 1 \end{bmatrix}, \qquad X = \begin{bmatrix} -6 & 2 \\ 2 & -5 \end{bmatrix} \qquad (5.106)$$

Case 4. $R^2 + R - 2I$

$$X = \begin{bmatrix} 2 & -2 \\ -2 & -3 \end{bmatrix}, \qquad X = \begin{bmatrix} -2 & -2 \\ -2 & -1 \end{bmatrix} \qquad (5.107)$$

Case 5. $R^2 + 2R + I$ - No solutions

Case 6. $R^2 - R - 2I$

$$X = \begin{bmatrix} -2 & -2 \\ -2 & -1 \end{bmatrix}, \qquad X = \begin{bmatrix} -2 & -2 \\ -2 & -3 \end{bmatrix} \qquad (5.108)$$

Case 7. $R^2 + 3R + 2I$

$$X = \begin{bmatrix} 6 & 2 \\ 2 & 1 \end{bmatrix}, \qquad X = \begin{bmatrix} -6 & 2 \\ 2 & -5 \end{bmatrix} \qquad (5.109)$$

Case 8. $R^2 - 2R + 4I$ - No solutions

Case 9. $R^2 - 4I$

$$X = \begin{bmatrix} 0 & -1 \\ -4 & -2 \end{bmatrix}, \qquad X = \begin{bmatrix} 0 & -4 \\ -1 & -2 \end{bmatrix} \qquad (5.110)$$

Case 10. $R^2 + 2R + 4I$ - No solutions

All solutions can be shown to satisfy Eq (5.12).

Example 5.3. The techniques of this chapter can also be used to solve the Liapunov equation by setting D equal to the zero matrix. The following example is taken from Hagander (Ref 31).

$$A = \begin{bmatrix} -3 & 0 \\ 0 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{5.111}$$

$$R = \begin{bmatrix} -B & D \\ -C & A \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & -3 & 0 \\ 0 & -1 & 0 & -2 \end{bmatrix} \tag{5.112}$$

$$|R-\lambda I| = (\lambda+3)(\lambda-3)(\lambda+2)(\lambda-2) = 0 \tag{5.113}$$

TABLE III

Combinations of Characteristic Roots of R

	2	-2	3	-3
2	$(\lambda-2)^2$	$(\lambda-2)(\lambda+2)$	$(\lambda-2)(\lambda-3)$	$(\lambda-2)(\lambda+3)$
-2	$(\lambda+2)(\lambda-2)$	$(\lambda+2)^2$	$(\lambda+2)(\lambda-3)$	$(\lambda+2)(\lambda+3)$
3	$(\lambda-3)(\lambda-2)$	$(\lambda-3)(\lambda+2)$	$(\lambda-3)^2$	$(\lambda-3)(\lambda+3)$
-3	$(\lambda+3)(\lambda-2)$	$(\lambda+3)(\lambda+2)$	$(\lambda+3)(\lambda-3)$	$(\lambda+3)^2$

Table III shows there are ten cases to consider, however, there are solutions for only two cases in this example:

Case 1. $R^2 + 5R + 6I$

$$\begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} + \begin{bmatrix} 15 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ -5 & 0 & -15 & 0 \\ 0 & -5 & 0 & -10 \end{bmatrix} + \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 30 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \end{bmatrix} = \begin{bmatrix} U & M \\ V & N \end{bmatrix} \tag{5.114}$$

$$X = -VU^{-1} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1/30 & 0 \\ 0 & 1/20 \end{bmatrix} = \begin{bmatrix} 1/6 & 0 \\ 0 & 1/4 \end{bmatrix} \tag{5.115}$$

Case 2. $R^2 - 5R + 6I$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 5 & 0 & 30 & 0 \\ 0 & 5 & 0 & 20 \end{bmatrix} = \begin{bmatrix} U & M \\ V & N \end{bmatrix} \tag{5.116}$$

$$X = N^{-1}V = \begin{bmatrix} 1/30 & 0 \\ 0 & 1/20 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 1/6 & 0 \\ 0 & 1/4 \end{bmatrix} \tag{5.117}$$

These solutions can be shown to satisfy Eq (2.1). Liapunov equations will always have D as a zero matrix which will always produce a zero matrix M. Therefore, as illustrated above, the only two possible combinations which produce solutions are $X = -VU^{-1}$ and $X = N^{-1}V$.

5.4 A Possible Iterative Solution Technique

In this section a complex-valued function f(z) of a complex variable z defined on the spectrum of $\tilde{R}$ is used as in P. Lancaster (Ref 32). The following form of Eq (5.12) is used in this section:

$$AX + XB = C + XDX \tag{5.118}$$

Theorem 5.14. If $f(z) = (z + a)(z - a)^{-1}$ where $a \neq 0$ and a is real, and if

$$f(A) \triangleq (aI - A)^{-1}(aI + A) = U \tag{5.119}$$

$$f(B) \triangleq (aI + B)(aI - B)^{-1} = V \tag{5.120}$$

where U and V are defined as in P. Lancaster (Ref 32) such that $(aI - A)^{-1}$ and $(aI - B)^{-1}$ exist, then a solution of

$$X = UXV - \frac{1}{2a}(U - I)C(V - I) - \frac{1}{2a}(U - I)XDX(V - I) \tag{5.121}$$

is also a solution of Eq (5.118).

Proof. Manipulation of Eq (5.118) gives

$$AX + XB = C + XDX$$

$$2aAX + 2aXB = 2aC + 2aXDX, \text{ (where } a \neq 0 \text{ and } a \text{ is real)}$$

$$a^2X + aXB + aAX + AXB - a^2X + aXB + aAX - AXB = 2aC + 2aXDX \tag{5.122}$$

$$aX(aI + B) + AX(aI + B) - aX(aI - B) + AX(aI - B) = 2aC + 2aXDX$$

$$(aI + A)X(aI + B) - (aI - A)X(aI - B) = 2aC + 2aXDX$$

If a is chosen so that $(aI - A)^{-1}$ and $(aI - B)^{-1}$ exist, then multiplying on the left and right gives

$$\begin{aligned}(aI - A)^{-1}(aI + A)X(aI + B)(aI - B)^{-1} - \\ (aI - A)^{-1}(aI - A)X(aI - B)(aI - B)^{-1} \\ = 2a(aI - A)^{-1}C(aI - B)^{-1} + 2a(aI - A)^{-1}XDX(aI - B)^{-1}\end{aligned} \tag{5.123}$$

Using Eqs (5.119) and (5.120)

$$UXV - X = 2a(aI - A)^{-1}C(aI - B)^{-1} + 2a(aI - A)^{-1}XDX(aI - B)^{-1} \tag{5.124}$$

$$f(z) = \frac{(z + a)}{(z - a)} \rightarrow f(z) - 1 = \frac{(z + a)}{(z - a)} \frac{-1}{1} = \frac{z + a - z + a}{z - a} =$$

$$2a(z - a)^{-1} \rightarrow f(A) - I = 2a(A - aI)^{-1} \tag{5.125}$$

$$U - I = 2a(A - aI)^{-1} \rightarrow \frac{1}{2a}(U - I) = (A - aI)^{-1} \tag{5.126}$$

$$\frac{1}{2a}(V - I) = (B - aI)^{-1} \tag{5.127}$$

$$X - UXV = -2a(aI - A)^{-1}C(aI - B)^{-1} - 2a(aI - A)^{-1}XDX(aI - B)^{-1} \tag{5.128}$$

where
$$(A - aI)^{-1} = -(aI - A)^{-1} \tag{5.129}$$

$$X - UXV = (U - I)C[-\frac{1}{2a}(V - I)] + (U - I)XDX[-\frac{1}{2a}(V - I)] \tag{5.130}$$

$$X - UXV = -\frac{1}{2a}(U - I)C(V - I) - \frac{1}{2a}(U - I)XDX(V - I) \tag{5.131}$$

$$X = UXV - \frac{1}{2a}(U - I)C(V - I) - \frac{1}{2a}(U - I)XDX(V - I) \tag{5.132}$$

If the conditions of Theorem 5.14 hold, then Eq (5.121) can be used to obtain a recursion formula:

$$X_{n+1} = UX_nV - \frac{1}{2a}(U - I)C(V - I) - \frac{1}{2a}(U - I)X_nDX_n(V - I) \tag{5.133}$$

which can be used to find the solution of Eq (5.118) using a suitable norm as an initial guess. For a good initial choice of X_0 the following example shows that by using Eq (5.133) it may be possible to develop a useful algorithm for computing the solution of the original Riccati equation.

Example 5.4. Another example from Martensson (Ref 28:44) is used to illustrate the recursion formula of Eq (5.133). Rewriting Martensson's equation in the form of Eq (5.118) gives:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} X + X \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + X \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} X \tag{5.134}$$

$$|A - \lambda I| = \lambda^2 - 1 = 0 \tag{5.135}$$

and

$$\lambda = \pm 1 \tag{5.136}$$

Choosing a = 0.1 yields the following:

$$U - I = 2a(A - aI)^{-1} = .2\begin{bmatrix} -.1 & 1 \\ 1 & -.1 \end{bmatrix}^{-1} = \frac{2}{9.9}\begin{bmatrix} .1 & 1 \\ 1 & .1 \end{bmatrix} \tag{5.137}$$

and

$$V - I = \frac{2}{9.9}\begin{bmatrix} .1 & 1 \\ 1 & .1 \end{bmatrix} \tag{5.138}$$

For an initial guess of

$$X_0 = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \tag{5.139}$$

Eq (5.133) produces

$$\begin{aligned} X_1 &= \begin{bmatrix} 5.07 & 3.40 \\ 3.40 & 3.07 \end{bmatrix} + \begin{bmatrix} .17 & -.17 \\ -.17 & .17 \end{bmatrix} + \begin{bmatrix} -.99 & -.99 \\ -.99 & -.99 \end{bmatrix} \\ &= \begin{bmatrix} 4.25 & 2.24 \\ 2.24 & 2.28 \end{bmatrix} \end{aligned} \tag{5.140}$$

and

$$X_2 = \begin{bmatrix} 4.36 & 2.36 \\ 2.36 & 2.42 \end{bmatrix} \tag{5.141}$$

After two iterations X compares favorably to Martensson's solution:

$$X = \begin{bmatrix} 4.41 & 2.41 \\ 2.41 & 2.41 \end{bmatrix} \tag{5.142}$$

Further investigation of Eq (5.121) should yield more specific results.

VI. Solutions of the Generalized Matrix Riccati Equation

The generalized form of the matrix Riccati equation arising in transport theory, as given in a series of papers by Bellman and Vasudevan (Ref 18; Ref 19), is considered in this chapter. The usual matrix Riccati equation can be related to the transport of a beam of particles through a medium of finite length with conservative type of interactions. By varying the picture of the interactions in the medium, the above authors arrive at extensions of the Riccati equation with higher order nonlinear terms. They made use of a linearization of these generalized equations.

The form of the Riccati equation

$$AX - XB = C + XDX + XEXFX \tag{6.1}$$

treated in this chapter can be derived as follows:

If one considers a system of linear vector equations

$$-Z' = AZ + DXEY \tag{6.2}$$

and

$$Y' = BZ + CY \tag{6.3}$$

with boundary conditions

$$\begin{aligned} Z(a) &= C \\ Y(0) &= 0 \end{aligned} \tag{6.4}$$

then, from the system, Eqs (6.2) and (6.3), the goal is finding Y(a), namely

$$Y(a) = X(a)Z(a) \tag{6.5}$$

where Z(a) is nonsingular.

First, a differential equation involving X(a) is obtained by differentiating Eq (6.5) with respect to a and making use of Eqs (6.2) and (6.3):

$$\begin{aligned} Y' &= X'Z + XZ' \\ Y' &= X'Z + X(-AZ - DXEY) \\ Y' &= X'Z - XAZ - XDXEY \\ BZ + CY &= X'Z - XAZ - XDXEY \\ BZ + CXZ &= X'Z - XAZ - XDXEXZ \end{aligned} \tag{6.6}$$

therefore, $$X' = B + CX + XA + XDXEX$$

The last of Eqs (6.6) can be written as

$$X' = B + CX + XA + XDX + XEXFX \tag{6.7}$$

In this chapter sufficient conditions for the existence of a solution of the nonlinear equation associated with the last of Eqs (6.6) are established.

6.1. A Necessary Condition for the Existence of Solutions of Eq (6.1)

Theorem 6.1. If X is a solution of the matrix differential equation

$$\frac{dX}{dt} + XB - AX + XDX + XEXFX + C = 0 \tag{6.8}$$

then the following matrices

$$\begin{bmatrix} A & \frac{dX}{dt} + XDX + XEXFX + C \\ 0 & B \end{bmatrix}, \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \tag{6.9}$$

are similar.

Proof. If X is a solution of Eq (6.8), then

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & \frac{dX}{dt} + XDX + XEXFX + C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} A & \frac{dX}{dt} + XDX + XEXFX + C + XB \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \tag{6.10}$$

$$= \begin{bmatrix} A & -AX + \frac{dX}{dt} + XDX + XEXFX + C + XB \\ 0 & B \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

and the matrices of Eq (6.9) are seen to be similar.

6.2. Sufficient Conditions for the Existence of Solutions of Eq (6.1)

In this section four theorems concerning sufficient conditions for solutions to the matrix differential equation and the algebraic equation are established. These results extend those of W. Roth (Ref 27) and R. Bellman (Ref 18; Ref 19).

Theorem 6.2. A sufficient condition for the algebraic matrix Riccati equation

$$AX - XB = C + XDXEX \tag{6.11}$$

to have a solution is that the equation

$$[A - \lambda I]X(\lambda) - Y(\lambda)[B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} \tag{6.12}$$

have a solution $X(\lambda)$, $Y(\lambda)$ where

$$X(\lambda) = X_0 + \lambda X_1 + \lambda^2 X_2 + \cdots + \lambda^p X_p \tag{6.13}$$

$$Y(\lambda) = Y_0 + \lambda Y_1 + \lambda^2 Y_2 + \cdots + \lambda^p Y_p \tag{6.14}$$

and

$$\hat{X} = X_0 + X_1 B + X_2 B^2 + \cdots + X_p B^p \tag{6.15}$$

Proof. Rewriting Eq (6.12) using Eqs (6.13) and (6.14) gives the following equation:

$$[A - \lambda I][X_0 + \lambda X_1 + \lambda^2 X_2 + \cdots + \lambda^p X_p] -$$
$$[Y_0 + \lambda Y_1 + \lambda^2 Y_2 + \cdots + \lambda^p Y_p][B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} \qquad (6.16)$$

Equating coefficients of like powers of λ and multiplying the equations on the right by $I, B, B^2, \ldots, B^{p+1}$ gives the following set of equations:

$$\begin{cases} AX_0 - Y_0B = C + \hat{X}D\hat{X}E\hat{X} \\ AX_1B - X_0B - Y_1B^2 + Y_0B = 0 \\ AX_2B^2 - X_1B^2 - Y_2B^3 + Y_1B^2 = 0 \\ AX_3B^3 - X_2B^3 - Y_3B^4 + Y_2B^3 = 0 \\ \quad\vdots \\ AX_{p-1}B^{p-1} - X_{p-2}B^{p-1} - Y_{p-1}B^p + Y_{p-2}B^{p-1} = 0 \\ AX_pB^p - X_{p-1}B^p - Y_pB^{p+1} + Y_{p-1}B^p = 0 \\ -X_pB^{p+1} + Y_pB^{p+1} = 0 \end{cases} \qquad (6.17)$$

Columnwise addition gives

$$[AX_0 + AX_1B^2 + AX_2B^3 + \cdots + AX_{p-1}B^{p-1} + AX_pB^p] -$$
$$[X_0B + X_1B^2 + X_2B^3 + \cdots + X_{p-1}B^p + X_pB^{p+1}] = C + \hat{X}D\hat{X}E\hat{X} \qquad (6.18)$$

Finally,

$$A[X_0 + X_1B + X_2B^2 + \cdots + X_{p-1}B^{p-1} + X_pB^p] -$$
$$[X_0 + X_1B + X_2B^2 + \cdots + X_{p-1}B^{p-1} + X_pB^p]B = C + \hat{X}D\hat{X}E\hat{X} \quad (6.19)$$

and so

$$A\hat{X} - \hat{X}B = C + \hat{X}D\hat{X}E\hat{X} \quad (6.20)$$

holds.

Theorem 6.3. A sufficient condition for Eq (6.11) to have a solution is that the equation

$$[A - \lambda I]X(\lambda) - Y(\lambda)[B - \lambda I] = C + \hat{Y}D\hat{Y}E\hat{Y} \quad (6.21)$$

have solutions $X(\lambda)$, $Y(\lambda)$ where Eqs (6.13), (6.14) and

$$\hat{Y} = Y_0 + AY_1 + A^2Y_2 + \cdots + A^pY_p \quad (6.22)$$

hold.

Proof. Rewriting Eq (6.21) using Eqs (6.13) and (6.14) produces the following equation:

$$[A - \lambda I][X_0 + \lambda X_1 + \lambda^2X_2 + \cdots + \lambda^pX_p] -$$
$$[Y_0 + \lambda Y + \lambda^2Y_2 + \cdots + \lambda^pY_p][B - \lambda I] = C + \hat{Y}D\hat{Y}E\hat{Y} \quad (6.23)$$

Equating coefficients of like powers of λ and multiplying the equations on the left by $I, A, A^2, \ldots, A^{p+1}$ gives the following set of equations:

$$
\begin{cases}
AX_0 & & - Y_0B & & = C + \hat{Y}D\hat{Y}E\hat{Y} \\
A^2X_1 & - AX_0 & - AY_1B & + AY_0 & = 0 \\
A^3X_2 & - A^2X_1 & - A^2Y_2B & + A^2Y_1 & = 0 \\
A^4X_3 & - A^3X_2 & - A^3Y_3B & + A^3Y_2 & = 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
A^pX_{p-1} & - A^{p-1}X_{p-2} & - A^{p-1}Y_{p-1}B & + A^{p-1}Y_{p-2} & = 0 \\
A^{p+1}X_p & - A^pX_{p-1} & - A^pY_pB & + A^pY_{p-1} & = 0 \\
 & - A^{p+1}X_p & & + A^{p+1}Y_p & = 0
\end{cases}
\tag{6.24}
$$

Columnwise addition gives

$$
[AY_0 + A^2Y_1 + A^3Y_2 + \cdots + A^pY_{p-1} + A^{p+1}Y_p] -
$$

$$
[Y_0B + AY_1B + A^2Y_2B + \cdots + A^{p-1}Y_{p-1}B + A^pY_pB] = C + \hat{Y}D\hat{Y}E\hat{Y} \tag{6.25}
$$

Finally,

$$
A[Y_0 + AY_1 + A^2Y_2 + \cdots + A^{p-1}Y_{p-1} + A^pY_p] -
$$

$$
[Y_0 + AY_1 + A^2Y_2 + \cdots + A^{p-1}Y_{p-1} + A^pY_p]B = C + \hat{Y}D\hat{Y}E\hat{Y} \tag{6.26}
$$

and so

$$
A\hat{Y} - \hat{Y}B = C + \hat{Y}D\hat{Y}E\hat{Y} \tag{6.27}
$$

holds.

Example 6.1. An illustration of the above theorem is given by the following one dimensional example:

$$2y = 1 + y^2 \tag{6.28}$$

which can be written as

$$[3][y] - [y][1] = [1] + [y][1][y] \tag{6.29}$$

where $A = 3, \quad B = 1 \quad C = 1, \quad D = 1$ (6.30)

The equation

$$[A - \lambda I]X(\lambda) - Y(\lambda)[B - \lambda I] = C + \hat{Y} \cdot D \cdot \hat{Y} \tag{6.31}$$

has solution $\hat{Y} = Y_0 = 1$ (6.32)

Simplifying the above results in the following:

$$\begin{aligned} [A - \lambda I]X(\lambda) - 1[B - \lambda I] &= 1 + 1 \cdot 1 \cdot 1 \\ [3 - \lambda]X(\lambda) - 1[1 - \lambda] &= 2 \\ [3 - \lambda]X(\lambda) &= 2 + 1 - \lambda \\ X(\lambda) &= \frac{3 - \lambda}{3 - \lambda} = 1 \end{aligned} \tag{6.33}$$

and 1 is easily seen to be the solution of Eq (6.28).

Theorem 6.4. *A sufficient condition that the matrix differential* Riccati Equation

$$AX - XB = C + XDXEX + \frac{dX}{dt} \tag{6.34}$$

have a solution is that the equation

$$[A - \lambda I]X(\lambda,t) - Y(\lambda,t)[B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \tag{6.35}$$

have a solution $X(\lambda,t)$, $Y(\lambda,t)$ where

$$X(\lambda,t) = X_0 + tX_1\lambda + t^2X_2\lambda^2 + t^3X_3\lambda^3 + \cdots \tag{6.36}$$

$$Y(\lambda,t) = Y_0 + tY_1\lambda + t^2Y_2\lambda^2 + t^3Y_3\lambda^3 + \cdots \tag{6.37}$$

and

$$\hat{X}(t) = X_0 + X_1Bt + X_2B^2t^2 + X_3B^3t^3 + \cdots \tag{6.38}$$

Proof. Rewriting Eq (6.35) using Eqs (6.36) and (6.37) produces the following equation:

$$[A - \lambda I][X_0 + tX_1\lambda + t^2X_2\lambda^2 + \cdots] -$$
$$[Y_0 + tY_1\lambda + t^2Y_2\lambda^2 + \cdots][B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \tag{6.39}$$

Equating coefficients of like powers of λ and multiplying the following equations on the right by I, B, B^2, . . . , gives the following set of equations:

$$\begin{cases}
AX_0 - Y_0B = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \\
AX_1tB - X_0B - Y_1B^2t + Y_0B = 0 \\
AX_2t^2B^2 - X_1tB^2 - Y_2B^3t^2 + Y_1tB^2 = 0 \\
AX_3t^3B^3 - X_2t^2B^3 - Y_3B^4t^3 + Y_2t^2B^3 = 0 \\
\quad\vdots \\
AX_pt^pB^p - X_{p-1}t^{p-1}B^p - Y_pB^{p+1}t^p + Y_{p-1}t^{p-1}B^p = 0 \\
\qquad - X_pt^pB^{p+1} + Y_pt^pB^{p+1} = 0
\end{cases} \tag{6.40}$$

Columnwise addition gives

$$[AX_0 + AXtB + AX_2t^2B^2 + \cdots + AX_pt^pB^p + \cdots] -$$
$$[X_0B + X_1tB^2 + X_2t^2B^3 + \cdots + X_pt^pB^{p+1} + \cdots] = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \tag{6.41}$$

$$A[X_0 + X_1Bt + X_2B^2t^2 + \cdots + X_pB^pt^p + \cdots] -$$

$$[X_0 + X_1Bt + X_2B^2t^2 + \cdots + X_pB^pt^p + \cdots]B = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \quad (6.42)$$

and finally,

$$A\hat{X}(t) - \hat{X}(t)B = C + \hat{X}(t)D\hat{X}(t)E\hat{X}(t) + \frac{d\hat{X}(t)}{dt} \quad (6.43)$$

Example 6.2. The preceding theorem is illustrated by the following scalar Riccati differential equation;

$$[1]X(t) - X(t)[-1] = [1] + X(t)[1]X(t) + \frac{dX}{dt} \quad (6.44)$$

where $A = 1$, $B = -1$, $C = 1$, $D = 1$

The equation

$$[A - \lambda I]X(t,\lambda) - Y(t,\lambda)[B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} + \frac{d\hat{X}}{dt} \quad (6.45)$$

has a solution $X(\lambda,t)$, $Y(\lambda,t)$ given below:

$$\begin{aligned} Y(\lambda,t) &= 1 + \frac{1}{C}\lambda - \frac{t}{C^2}\lambda^2 + \frac{t^2}{C^3}\lambda^3 - \frac{t^3}{C^4}\lambda^4 + \frac{t^4}{C^5}\lambda^5 - \cdots \\ &= Y_0 + Y_1\lambda + Y_2\lambda^2 + Y_3\lambda^3 + \cdots \end{aligned} \quad (6.46)$$

$$\begin{aligned} X(\lambda,t) &= 1 + \frac{1}{C}\lambda - \frac{t}{C^2}\lambda^2 + \frac{t^2}{C^3}\lambda^3 - \frac{t^3}{C^4}\lambda^4 + \frac{t^4}{C^5}\lambda^5 - \cdots \\ &= X_0 + X_1\lambda + X_2\lambda^2 + X_3\lambda^3 + \cdots \end{aligned} \quad (6.47)$$

Then, by Eq (6.38),

$$\begin{aligned} \hat{X}(t) &= 1 + \frac{1}{C} - \frac{t}{C^2} + \frac{t^2}{C^3} - \frac{t^3}{C^4} + \frac{t^4}{C^5} - \frac{t^5}{C^6} + \cdots \\ &= X_0 + X_1Bt + X_2B^2t^2 + X_3B^3t^3 + \cdots \end{aligned} \quad (6.48)$$

and
$$\frac{d\hat{X}(t)}{dt} = -\frac{1}{c^2} + \frac{2t}{c^3} - \frac{3t^2}{c^4} + \frac{4t^3}{c^5} - \frac{5t^4}{c^6} + \frac{6t^5}{c^7} + \cdots$$
$$\frac{dX_0}{dt} + \frac{dX_1}{dt} Bt + \frac{dX_2}{dt} B^2t^2 + \frac{dX_3}{dt} B^3t^3 + \cdots \tag{6.49}$$

and Eq (6.44) has a solution $\hat{X}(t)$ given in Eqs (6.48) and (6.49) which can be written in the form

$$\hat{X}(t) = 1 + \frac{1}{(t + c)}, \qquad c = \text{real number} \neq 0 \tag{6.50}$$

Theorem 6.5. A sufficient condition that Eq (6.34) have a solution is that the equation

$$[A - \lambda I]X(\lambda,t) - Y(\lambda,t)[B - \lambda I] = C + \hat{Y}D\hat{Y}E\hat{Y} + \frac{d\hat{Y}}{dt} \tag{6.51}$$

have a solution $X(\lambda,t)$, $Y(\lambda,t)$ where Eqs (6.36), (6.37), and

$$\hat{Y}(t) = Y_0 + AY_1t + A^2Y_2t^2 + AY_3t^3 + \cdots \tag{6.52}$$

hold.

Proof. Rewriting Eq (6.51) using Eqs (6.36) and (6.37) produces the following equation:

$$[A - \lambda I][X_0 + tX_1\lambda + t^2X_2\lambda^2 + t^3X_3\lambda^3 + \cdots] -$$
$$[Y_0 + tY_1\lambda + t^2Y_2\lambda^2 + t^3Y_3\lambda^3 + \cdots][B - \lambda I] = C + \hat{Y}D\hat{Y}E\hat{Y} + \frac{d\hat{Y}}{dt} \tag{6.53}$$

Equating like powers of λ and multiplying the following equations on the left by I, A, A^2, . . . gives the following set of equations:

$$\begin{cases} AX_0 - Y_0B = C + \hat{Y}D\hat{Y}E\hat{Y} + \frac{d\hat{Y}}{dt} \\ A^2X_1 - AX_0 - AY_1B + AY_0 = 0 \\ A^3X_2 - A^2X_1 - A^2Y_2B + A^2Y_1 = 0 \end{cases}$$

$$\begin{cases} A^4X_3 - A^3X_2B - A^3Y_3B + A^3Y_2 = 0 \\ \quad\cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \\ \quad\cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \\ \quad\cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot \\ A^{p+1}X_p - A^pY_{p-1} - A^pY_pB + A^pY_{p-1} = 0 \\ \qquad\quad - A^{p+1}Y_p \qquad\quad + A^{p+1}Y_p = 0 \end{cases} \tag{6.54}$$

Columnwise addition gives:

$$[AY_0 + A^2Y_1t + A^3Y_2t^2 + \cdots + A^{p+1}Y_pt^p + \cdots] -$$
$$[Y_0B + AY_1tB + A^2Y_2t^2B + \cdots + A^pY_pt^pB + \cdots]$$
$$= C + \hat{Y}D\hat{Y}E\hat{Y} + \frac{d\hat{Y}}{dt} \tag{6.55}$$

$$A[Y_0 + AY_1t + A^2Y_2t^2 + \cdots + A^pY_pt^p + \cdots] -$$
$$[Y_0 + AY_1t + A^2Y_2t^2 + \cdots + A^pY_pt^p + \cdots]B = C + \hat{Y}D\hat{Y}E\hat{Y} + \frac{d\hat{Y}}{dt} \tag{6.56}$$

and finally,

$$A\hat{Y}(t) - \hat{Y}(t)B = C + \hat{Y}(t)D\hat{Y}(t)E\hat{Y}(t) + \frac{d\hat{Y}(t)}{dt} \tag{6.57}$$

The following three examples illustrate the techniques presented in Theorems 6.2 and 6.3 for the algebraic matrix Riccati equation. All three present solutions to the equation

$$AX - XB = C + XDXEX \tag{6.11}$$

using increasing powers of λ in the intermediate solution matrices.

Example 6.3. If

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & -3 \\ -2 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.58}$$

then the following exist:

$$\begin{aligned} X(\lambda) &= \begin{bmatrix} \lambda+2 & 3 \\ 3 & \lambda+2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = X_0 + \lambda X_1 \\ Y(\lambda) &= \begin{bmatrix} \lambda-1 & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = Y_0 + \lambda Y_1 \end{aligned} \tag{6.59}$$

Now

$$\begin{aligned} \hat{X} &= X_0 + X_1 B = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -3 \\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ \hat{Y} &= Y_0 + AY_1 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \end{aligned} \tag{6.60}$$

and the following equation is satisfied:

$$[A - \lambda I]X(\lambda) - Y(\lambda)[B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} \tag{6.12}$$

$$\begin{aligned} &\begin{bmatrix} 1-\lambda & 0 \\ 1 & -\lambda \end{bmatrix} \begin{bmatrix} \lambda+2 & 3 \\ 3 & \lambda+2 \end{bmatrix} - \begin{bmatrix} \lambda-1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} -2-\lambda & -3 \\ -2 & -2-\lambda \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ &\begin{bmatrix} -\lambda^2-\lambda+2 & 3-3\lambda \\ -2\lambda+2 & 3-2\lambda-\lambda^2 \end{bmatrix} - \begin{bmatrix} -\lambda^2-\lambda+2 & -3\lambda+3 \\ -2\lambda & -2\lambda-\lambda^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &\begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix} \end{aligned} \tag{6.61}$$

Also, $\hat{X}$ satisfies

$$A\hat{X} - \hat{X}B = C + \hat{X}D\hat{X}E\hat{X} \tag{6.20}$$

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} - \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}-2 & -3\\-2 & -2\end{bmatrix} = \begin{bmatrix}0 & 0\\2 & 3\end{bmatrix} +$$
$$\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} \tag{6.62}$$
$$\begin{bmatrix}0 & 0\\2 & 3\end{bmatrix} = \begin{bmatrix}0 & 0\\2 & 3\end{bmatrix}$$

while $\hat{Y}$ satisfies

$$A\hat{Y} - \hat{Y}B = C + \hat{Y}D\hat{Y}E\hat{Y} \tag{6.27}$$

Example 6.4. If

$$A = \begin{bmatrix}1 & 1\\0 & 0\end{bmatrix},\quad B = \begin{bmatrix}-1 & -1\\-1 & -2\end{bmatrix},\quad C = \begin{bmatrix}-35 & -71\\-117 & -237\end{bmatrix},\quad D = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix},\quad E = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \tag{6.63}$$

then the following exist:

$$X(\lambda) = \begin{bmatrix}(\lambda+1)^2 & 1\\0 & \lambda^2+1\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} + \lambda\begin{bmatrix}2 & 0\\0 & 0\end{bmatrix} +$$
$$\lambda^2\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = X_0 + \lambda X_1 + \lambda^2 X_2$$
$$Y(\lambda) = \begin{bmatrix}\lambda^2+1 & -2\lambda+5\\-\lambda+3 & \lambda^2-2\lambda+6\end{bmatrix} = \begin{bmatrix}1 & 5\\3 & 6\end{bmatrix} + \lambda\begin{bmatrix}0 & -2\\-1 & -2\end{bmatrix} + \tag{6.64}$$
$$\lambda^2\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = Y_0 + \lambda Y_1 + \lambda^2 Y_2$$

Now

$$\hat{X} = X_0 + X_1 B + X_2 B^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix} +$$

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

$$\hat{Y} = Y_0 + AY_1 + A^2 Y_2 = \begin{bmatrix} 1 & 5 \\ 3 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -2 \\ -1 & -2 \end{bmatrix} + \tag{6.65}$$

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

and the following equation is satisfied:

$$[A - \lambda I]X(\lambda) - Y(\lambda)[B - \lambda I] = C + \hat{X}D\hat{X}E\hat{X} \tag{6.12}$$

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & -\lambda \end{bmatrix} \begin{bmatrix} (\lambda+1)^2 & 1 \\ 0 & \lambda^2+1 \end{bmatrix} - \begin{bmatrix} \lambda^2+1 & -2\lambda+5 \\ -\lambda+3 & \lambda^2-2\lambda+6 \end{bmatrix} \begin{bmatrix} -1-\lambda & -1 \\ -1 & -2-\lambda \end{bmatrix}$$

$$= \begin{bmatrix} -35 & -71 \\ -117 & -237 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \tag{6.60}$$

$$\begin{bmatrix} 7 & 13 \\ 9 & 15 \end{bmatrix} = \begin{bmatrix} 7 & 13 \\ 9 & 15 \end{bmatrix}$$

Also, $\hat{X}$ satisfies Eq (6.20):

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

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{6.67}$$

$$\begin{bmatrix} 7 & 13 \\ 9 & 15 \end{bmatrix} = \begin{bmatrix} 7 & 13 \\ 9 & 15 \end{bmatrix}$$

while $\hat{Y}$ satisfies Eq (6.27).

Example 6.5. If

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 36 & -66 \\ -22 & -3 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.68}$$

then the following exist:

$$\begin{aligned} X(\lambda) &= \begin{bmatrix} \lambda^3+2\lambda+1 & 1 \\ 0 & \lambda^2+1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \lambda^3 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= X_0 + \lambda X_1 + \lambda^2 X_2 + \lambda^3 X_3 \\ Y(\lambda) &= \begin{bmatrix} \lambda^3-2\lambda^2+4\lambda-5 & -\lambda^2+4\lambda-11 \\ -\lambda^2+\lambda-3 & \lambda^2-\lambda+2 \end{bmatrix} \\ &= \begin{bmatrix} -5 & -11 \\ -3 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 4 & 4 \\ 1 & -1 \end{bmatrix} + \lambda^2 \begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix} + \lambda^3 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= Y_0 + \lambda Y_1 + \lambda^2 Y_2 + \lambda^3 Y_3 \end{aligned} \tag{6.69}$$

Now

$$\begin{aligned} \hat{X} &= X_0 + X_1 B + X_2 B^2 + X_3 B^3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix} + \\ &\quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & -7 \\ 0 & -8 \end{bmatrix} = \begin{bmatrix} -2 & -8 \\ 0 & 5 \end{bmatrix} \\ \hat{Y} &= Y_0 + AY_1 + A^2 Y_2 + A^3 Y_3 = \begin{bmatrix} -5 & -11 \\ -3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ 1 & -1 \end{bmatrix} + \\ &\quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -8 \\ 0 & 5 \end{bmatrix} \end{aligned} \tag{6.70}$$

and Eq (6.12) is satisfied:

$$\begin{bmatrix}1-\lambda & 0\\ 1 & -\lambda\end{bmatrix}\begin{bmatrix}\lambda^3+2\lambda+1 & 1\\ 0 & \lambda^2+1\end{bmatrix} - \begin{bmatrix}\lambda^3-2\lambda^2+4\lambda-5 & -\lambda^2+4\lambda-11\\ -\lambda^2+\lambda-3 & \lambda^2-\lambda+2\end{bmatrix}\begin{bmatrix}-1-\lambda & -1\\ 0 & -2-\lambda\end{bmatrix}$$

$$= \begin{bmatrix}36 & -66\\ -22 & -3\end{bmatrix} + \begin{bmatrix}-2 & -8\\ 0 & 5\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}-2 & -8\\ 0 & 5\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-2 & -8\\ 0 & 5\end{bmatrix} \qquad (6.71)$$

$$\begin{bmatrix}-4 & -26\\ -2 & -2\end{bmatrix} = \begin{bmatrix}-4 & -26\\ -2 & -2\end{bmatrix}$$

Also, $\hat{X}$ satisfies Eq (6.20):

$$\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}-2 & -8\\ 0 & 5\end{bmatrix} - \begin{bmatrix}-2 & -8\\ 0 & 5\end{bmatrix}\begin{bmatrix}-1 & -1\\ 0 & -2\end{bmatrix} = \begin{bmatrix}36 & -66\\ -22 & -3\end{bmatrix} + \begin{bmatrix}-40 & 40\\ 20 & 5\end{bmatrix} \qquad (6.72)$$

$$\begin{bmatrix}-4 & -26\\ -2 & 2\end{bmatrix} = \begin{bmatrix}-4 & -26\\ -2 & 2\end{bmatrix}$$

while $\hat{Y}$ satisfies Eq (6.27).

Example 6.6. This example demonstrates the techniques in Theorem 6.4 for the matrix differential equation

$$AX - XB = C + XDX - \frac{dX}{dt} \qquad (6.73)$$

where $A = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, $B = \begin{bmatrix}1 & 1\\ 0 & -1\end{bmatrix}$, $C = \begin{bmatrix}-1 & 0\\ 2-2t-2t^2 & 1-2t-t^2\end{bmatrix}$, $D = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ (6.74)

The following are solutions of Eqs (6.36) and (6.37):

$$X(\lambda,t) = \begin{bmatrix}1 & 1\\ 2t\lambda & t\lambda\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} + \lambda\begin{bmatrix}0 & 0\\ 2t & t\end{bmatrix} = X_0(t) + \lambda X_1(t)$$

$$Y(\lambda,t) = \begin{bmatrix}1-2t & 1-t\\ 2t\lambda & t\lambda\end{bmatrix} = \begin{bmatrix}1-2t & 1-t\\ 0 & 0\end{bmatrix} + \lambda\begin{bmatrix}0 & 0\\ 2t & t\end{bmatrix} = Y_0(t) + \lambda Y_1(t) \qquad (6.75)$$

Now

$$\hat{X} = X_0 + X_1 B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2t & t \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2t & t \end{bmatrix}$$

$$\hat{Y} = Y_0 + AY_1 = \begin{bmatrix} 1-2t & 1-t \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 2t & t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2t & t \end{bmatrix} \tag{6.76}$$

and the following equation is satisfied:

$$[A - \lambda I]X(\lambda,t) - Y(\lambda,t)[B - \lambda I] = C + \hat{X}D\hat{X} - \frac{d\hat{X}}{dt} \tag{6.77}$$

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2t\lambda & t\lambda \end{bmatrix} - \begin{bmatrix} 1-2t & 1-t \\ 2t\lambda & t\lambda \end{bmatrix} \begin{bmatrix} 1-\lambda & 1 \\ 0 & -1-\lambda \end{bmatrix}$$

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

$$\begin{bmatrix} 1-\lambda-2t\lambda & 1-\lambda+t\lambda \\ 2t\lambda-2t\lambda^2 & t\lambda-t\lambda^2 \end{bmatrix} - \begin{bmatrix} 1-2t-\lambda+2t\lambda & -t-\lambda+t\lambda \\ 2t\lambda-2t\lambda^2 & t\lambda-t\lambda^2 \end{bmatrix} \tag{6.78}$$

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

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

Also, $\hat{X}$ satisfies

$$A\hat{X} - \hat{X}B = C + \hat{X}D\hat{X} - \frac{d\hat{X}}{dt} \tag{6.79}$$

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

$$= \begin{bmatrix} -1 & 0 \\ 2-2t-2t^2 & 1-2t-t^2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 2t & t \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2t & t \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix} \tag{6.80}$$

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

while $\hat{Y}$ satisfies

$$A\hat{Y} - \hat{Y}B = C + \hat{Y}D\hat{Y} - \frac{d\hat{Y}}{dt} \tag{6.81}$$

Equation (6.73) can replace Eq (6.34) for identical results in Theorem 6.4.

VII. A New Method for Computing Matrix Inverses

A generalization of the classical Newton-Raphson method of finding zeros of functions of a single variable is extended in this chapter to the case of finding zeros of functions of a matrix variable.

In the classical problem of finding zeros of a function f(x) by the Newton-Raphson method the following algorithm is used:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, 2, \ldots \tag{7.1}$$

to produce a sequence $\{x_{k+1}\}$ which converges to a zero (x*) of f(x) = 0. Care must be taken in choosing an initial x_0 so that the above sequence $\{x_{k+1}\}$ actually converges to a finite number, x*. If the algorithm is rewritten as

$$x_{k+1} = x_k(2 - ax_k), \qquad a = \text{real number}, \ a \neq 0 \tag{7.2}$$

then the sequence $\{x_{k+1}\}$ converges to $\frac{1}{a}$ if and only if x_0 is chosen so that $|1 - ax_0| < 1$ (Ref 32:30).

If f(x) is chosen as follows:

$$f(x) = x^{-1} - a \tag{7.3}$$

then

$$f'(x) = \frac{-1}{x^2} = -x^{-2} \tag{7.4}$$

and the classical Newton-Raphson algorithm may be written

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$= x_k - \frac{\frac{1}{x_k} - a}{-\frac{1}{x_k^2}}$$

$$= x_k + x_k^2 \left(\frac{1}{x_k} - a\right) \tag{7.5}$$

$$= x_k + x_k - x_k^2 a$$

$$= x_k(2 - ax_k)$$

If $y_k = 1 - ax_k$, then $y_k = y_0^{2^k}$ and it is easily seen that the condition $|1 - ax_0| < 1$ is a necessary and sufficient condition that $\{x_{k+1}\}$ converges to some x^* which is a zero of $f(x)$, that is, $f(x^*) = 0$.

Example 7.1. A one dimensional example illustrates the critical choice of x_0. To find $\frac{1}{a}$ given $a = 2$, $f(x) = \frac{1}{x} - a$, and $x_0 = 3$ the algorithm

$$x_{k+1} = x_k(2 - 2x_k) \tag{7.6}$$

generates the sequence {3, -12, -312, . . . } which diverges. However, if x_0 is chosen as 1/4 then the sequence {1/4, 3/8, 15/32, . . . } converges to 1/2, which is a zero of $f(x)$.

In the matrix case some conditions on the initial choice of X_0 are necessary to guarantee convergence of a sequence $\{X_{k+1}\}$ to X^*, a zero of a matrix function $F(X)$, that is, $F(X^*) = 0$.

Example 7.2. A matrix example illustrates the critical choice of X_0

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \qquad X_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.7}$$

If the following algorithm, which is established later, is used for the matrix case, namely,

$$X_{k+1} = X_k(2I - X_kA) \tag{7.8}$$

then
$$X_1 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = X_2 = X_3 = \cdots \tag{7.9}$$

However, for

$$X_0 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \tag{7.10}$$

then
$$X_1 = \begin{bmatrix} 0 & -4 \\ 0 & -3 \end{bmatrix} \tag{7.11}$$

$$X_2 = \begin{bmatrix} 0 & -20 \\ 0 & -15 \end{bmatrix} \tag{7.12}$$

$$X_3 = \begin{bmatrix} 0 & -600 \\ 0 & -450 \end{bmatrix} \tag{7.13}$$

In order to develop the algorithm for the matrix case the notion of the Fréchet derivative of a matrix function is required. X and Y denote Banach algebras which are complete, normed, linear vector spaces and $F(\cdot)$ is an operator from X into Y. If a bounded linear operator L from X into Y exists, such that

$$\lim_{\|\Delta X\| \to 0} \frac{\| F(X_0 + \Delta X) - F(X_0) - L(\Delta X) \|}{\| \Delta X \|} = 0 \tag{7.14}$$

then $F(\cdot)$ is said to be differentiable at X_0 and the bounded linear operator $F'(X_0) = L$ is called the first Fréchet derivative at $X = X_0$.

The following results are used to develop the algorithm for the matrix case given earlier, and to establish sufficient conditions on the

choice of X_0 to guarantee convergence of $\{X_{k+1}\}$ to some X* which is a zero of $F(X) = X^{-1} - A$.

Theorem 7.1. If $F(X) = X^{-1} - A$ is a mapping of a Banach algebra $\hat{X}$ into a Banach algebra $\hat{Y}$, where for a given $X \in \hat{X}$, $X^{-1} \in \hat{Y}$, and $(X^{-1} - A) \in \hat{Y}$, then the Fréchet derivative of $F(X) = X^{-1} - A$ is $F'(X) = -X^{-2}$.

Proof. Using the results of Lusternik and Sobolev (Ref 33:91), if X and ΔX are elements of a Banach algebra and if X has an inverse element X^{-1}, and $\| \Delta X \| < (\| X^{-1} \|)^{-1}$, then $(X + \Delta X)$ also has an inverse written as $(X + \Delta X)^{-1}$. If $\| \Delta X \| < (\| X^{-1} \|)^{-1}$ and if X^{-1} exists, then, letting $F(X) = X^{-1} - A$,

$$F(X + \Delta X) - F(X) = [(X + \Delta X)^{-1} - A] - [X^{-1} - A] \tag{7.15}$$

$$= (X + \Delta X)^{-1} - X^{-1} \tag{7.16}$$

If X^{-1} and Y^{-1} exist, then the following identity holds:

$$X^{-1}(X - Y)Y^{-1} = Y^{-1} - X^{-1} \tag{7.17}$$

Using Eqs (7.16) and (7.17), and letting $Y = X + \Delta X$,

$$\begin{aligned} F(X + \Delta X) - F(X) &= X^{-1}(X - X - \Delta X)(X + \Delta X)^{-1} \\ &= X^{-1}(-\Delta X)(X + \Delta X)^{-1} \\ &= -X^{-1}(\Delta X)(X + \Delta X)^{-1} \\ &= -X^{-1}(\Delta X)(X + \Delta X)^{-1} + X^{-1}(\Delta X)X^{-1} - X^{-1}(\Delta X)X^{-1} \end{aligned} \tag{7.18}$$

the following form of the above is desired:

$$F(X + \Delta X) - F(X) = F'(X)(\Delta X) + \alpha(X, \Delta X) \tag{7.19}$$

$$\lim_{\| \Delta X \| \to 0} \frac{\| \alpha(X, \Delta X) \|}{\| \Delta X \|} = 0 \tag{7.20}$$

Grouping the first two terms of the right side of the last of Eqs (7.18) gives

$$\begin{aligned} F(X + \Delta X) - F(X) &= -X^{-1}(\Delta X)[(X + \Delta X)^{-1} - X^{-1}] - X^{-1}(\Delta X)X^{-1} \\ &= -X^{-1}(\Delta X)X^{-1}(-\Delta X)(X + \Delta X)^{-1} - X^{-1}(\Delta X)X^{-1} \\ &= X^{-1}(\Delta X)X^{-1}(\Delta X)(X + \Delta X)^{-1} - X^{-1}(\Delta X)X^{-1} \end{aligned} \tag{7.21}$$

Now $$\frac{\| \alpha(X,\Delta X) \|}{\| \Delta X \|} = \frac{\| X^{-1} \|^2 \cdot \| \Delta X \|^2 \cdot \| (X + \Delta X)^{-1} \|}{\| \Delta X \|}$$

$$= \| X^{-1} \|^2 \cdot \| (X + \Delta X) \| \cdot \| \Delta X \| \to 0 \text{ as } \| \Delta X \| \to 0 \tag{7.22}$$

For ΔX sufficiently small, $(\Delta X)X^{-1} = X^{-1}(\Delta X)$ and the Fréchet derivative of $F(X)$ is given by

$$F'(X) = -X^{-2} = -(X^{-1})(X^{-1}) \tag{7.23}$$

Theorem 7.2 (Kantorovich). If for some $X_0 \in D_0$, $F'(X_0)$ exists, and

$$\begin{aligned} &\text{(i)} \quad \| [F'(X_0)]^{-1} \| \leq \beta \\ &\text{(ii)} \quad \| [F'(X_0)]^{-1} F(X_0) \| \leq \eta \\ &\text{(iii)} \quad \| F'(X) - F'(Y) \| \leq \kappa \| X - Y \|, \text{ for any } X, Y \in D_0 \end{aligned} \tag{7.24}$$

where $$h = \beta\eta\kappa \leq 1/2 \tag{7.25}$$

and $$\Omega_* = \left\{ X: \quad \| X - X_0 \| \leq t^* = \frac{(1-\sqrt{1 - 2h})}{h} \eta \right\} \tag{7.26}$$

holds. Now if $\Omega_* \in D_0$, then the Newton iterations

$$X_{k+1} = X_k - [F'(X_k)]^{-1} F(X_k) \tag{7.27}$$

are well defined, remain in Ω_*, and converge to $X^* \in \Omega_*$ such that $F(X^*) = 0$. In addition

$$\| X^* - X_k \| < \frac{\eta}{h} \cdot \frac{(1-\sqrt{1-2h})^{2^k}}{2^k}, \qquad k = 0, 1, 2, \ldots \tag{7.28}$$

Proof. A proof is given by Tapia (Ref 34:389-392).

Theorem 7.3. If F(X) is a matrix function satisfying the hypotheses of Theorems 7.1 and 7.2, then the inverse of a n by n nonsingular matrix A is given by the limit of the sequence

$$X_{k+1} = X_k - [F'(X_k)]^{-1} F(X_k), \qquad k = 0, 1, 2, \ldots \tag{7.29}$$

or

$$X_{k+1} = X_k [2I - X_k A], \qquad k = 0, 1, 2, \ldots \tag{7.30}$$

Proof. By choosing F(X) to be $X^{-1} - A$, the recursion formula follows from Eq (7.27):

$$\begin{aligned} X_{k+1} &= X_k - [F'(X_k)]^{-1} F(X_k) \\ &= X_k - [-X^{-2}]^{-1} [X^{-1} - A] \\ &= X_k + X_k - X_k^2 A \\ &= X_k [2I - X_k A] \end{aligned} \tag{7.31}$$

Example 7.3. To illustrate the formula just presented the following matrices are used:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad X_0 = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.32}$$

Now

$$X_{k+1} = X_k (2I - X_k A) \tag{7.30}$$

where $F(X) = -X^{-2} = -(X^{-1})(X^{-1})$ and $\| X_0 \| = \max |a_{ij}|$, i,j = 1, 2, . . . , n, and Eqs (7.24) become

$$(i)\ \| [F'(X_0)]^{-1} \| = \left\| \begin{bmatrix} -1/4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \right\| = \beta = 1 \tag{7.33}$$

$$(ii)\ \| [F'(X_0)]^{-1} F(X_0) \| = \| -[X_0^{-2}]^{-1} \|$$

$$= \| -X_0^{2}(X_0^{-1} - A) \|$$

$$= \| (-X_0 + X^2 A) \|$$

$$= \left\| \begin{bmatrix} -1/2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\| \tag{7.34}$$

$$= \left\| \begin{bmatrix} -1/2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 1/4 & 1/4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1/4 & 1/4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\| = \eta = 1/4$$

$$\begin{aligned} (iii)\ &\| F'(X) - F'(Y) \| \\ &= \| -X^{-2} + Y^{-2} \| \\ &= \| Y^{-2} - X^{-2} \| \\ &\leq \| (Y^{-1}X^{-1} - X^{-2}) + (Y^{-2} - Y^{-1}X^{-1}) \| \\ &\leq \| (Y^{-1} - X^{-1})X^{-1} \| + \| Y^{-1}(Y^{-1} - X^{-1}) \| \\ &\leq \| (X^{-1} - Y^{-1})X^{-1} \| + \| Y^{-1}(X^{-1} - Y^{-1}) \| \\ &\leq \| Y^{-1}(Y - X)X^{-1} \cdot X^{-1} \| + \| Y^{-1}(Y^{-1})(Y - X)X^{-1} \| \\ &\leq \| Y^{-1} \| \cdot \| X^{-1} \|^2 \cdot \| Y - X \| + \| Y^{-1} \|^2 \cdot \| Y - X \| \cdot \| X^{-1} \| \\ &\leq \| [\| Y^{-1} \| \cdot \| X^{-1} \|^2 + \| Y^{-1} \|^2 \cdot \| X^{-1} \|] \cdot \| Y - X \| \\ &< \kappa \cdot \| Y - X \| \end{aligned} \tag{7.35}$$

Therefore, for

$$h = \frac{\kappa}{4} \leq 1/2 \tag{7.36}$$

$$\kappa \leq 2 \tag{7.37}$$

For this example, in order to satisfy the requirement that $h \leq 1/2$ it is sufficient that $\kappa \leq 2$.

After some computation

$$X_1 = \begin{bmatrix} .75 & -.25 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{7.38}$$

$$X_2 = \begin{bmatrix} .94 & -.63 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.39}$$

$$X_3 = \begin{bmatrix} .996 & -.918 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.40}$$

$$X_4 = \begin{bmatrix} 1.00 & -.84 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \simeq \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.41}$$

which approximates A^{-1} after four iterations. If X_0 is chosen as

$$X_0 = \begin{bmatrix} .1 & 0 & 0 \\ 0 & .1 & 0 \\ 0 & 0 & .1 \end{bmatrix} \tag{7.43}$$

then

$$X_1 = \begin{bmatrix} .19 & -.01 & 0 \\ 0 & .19 & 0 \\ 0 & 0 & .19 \end{bmatrix} \tag{7.44}$$

$$X_2 = \begin{bmatrix} .34 & -.05 & 0 \\ 0 & .34 & 0 \\ 0 & 0 & .34 \end{bmatrix} \tag{7.45}$$

$$X_3 = \begin{bmatrix} .57 & -.19 & 0 \\ 0 & .57 & 0 \\ 0 & 0 & .57 \end{bmatrix} \tag{7.46}$$

$$X_4 = \begin{bmatrix} .81 & -.77 & 0 \\ 0 & .81 & 0 \\ 0 & 0 & .81 \end{bmatrix} \tag{7.47}$$

Further work is necessary to determine proper criteria for termination of this algorithm.

VIII. Conclusion

The increasing use of matrix equations in the engineering sciences has stimulated a commensurate interest in the investigation of new ways to solve these equations. The original computational methods presented in this report can simplify calculation of solutions of certain matrix equations. These new techniques are easily adapted for use by the digital computer, in some cases eliminating the need for lengthy iteration processes. Several new representations of solutions can possibly be extended resulting in additional areas of their application. The iteration scheme presented in the last chapter could form the basis of another technique for finding solutions of the Liapunov and Riccati equations.

AD-A032 667 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 12/1
SOLUTIONS OF CERTAIN MATRIX EQUATIONS WITH APPLICATIONS TO ENGI--ETC(U)
DEC 76 J L LEUTHAUSER

UNCLASSIFIED GA/MA/76D-1 NL

2 OF 2
AD
A032667



END

DATE
FILMED
1-77

Bibliography

1. Meirovitch, L. Methods of Analytical Dynamics. New York: McGraw-Hill, Inc., 1970.

2. LaSalle, J. and S. Lefschetz. Stability by Liapunov's Direct Method. New York: Academic Press, Inc., 1961.

3. Ma, E. C. "A Finite Series Solution of the Matrix Equation AX - XB = C." SIAM Journal on Applied Mathematics, 14:490-495 (May 1966).

4. Blackburn, T. R., and D. T. Vaughn. "Application of Linear Optimal Control and Filtering Theory to the Saturn V Launch Vehicle." IEEE Transactions on Automatic Control, AC-16:799-806 (December 1971).

5. Howerton, R. D. "A New Solution of the Discrete Algebraic Riccati Equation." IEEE Transactions on Automatic Control, AC-19:90-92 (February 1974).

6. Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems." ASME Journal of Basic Engineering, 82:35-46 (March 1960).

7. Deyst, J. J. "A Derivation of the Optimum Continuous Linear Estimator for Systems with Correlated Measurement Noise." AIAA Journal, 7:2116-2119 (September 1969).

8. Sutherland, A. A., Jr. and A. Gelb. Application of the Kalman Filter to Aided Inertial Systems. NWC TP 4652. Naval Weapons Center, China Lake, California, August 1968.

9. Bryson, A. E., Jr. and Y-C. Ho. Applied Optimal Control. Washington, D. C.: Hemisphere Publishing Corporation, 1975.

10. Breakwell, J. V. "The Optimization of Trajectories." SIAM Journal, 7:215-247 (June 1959).

11. McReynolds, S. R. "The Successive Sweep Method and Dynamic Programming." Journal of Mathematical Analysis and Applications, 19: 565-598 (September 1967).

12. Bender, E. K. "Optimum Linear Preview Control with Application to Vehicle Suspension." ASME Journal of Basic Engineering, 90:213-221 (June 1968).

13. Karnopp, D. C. and A. K. Trikha. "Comparative Study of Optimatization Techniques for Shock and Vibration Isolation." ASME Journal of Engineering for Industry, 91:1128-1132 (November 1969).

14. Sevin, F. and W. D. Pilkey. Optimum Shock and Vibration Isolation, SVM-6. The Shock and Vibration Information. Washington, D. C.: United States Department of Defense, 1971.

15. Yang, J.-N. "Application of Optimal Control Theory to Civil Engineering Structures." ASCE Journal of Engineering Mechanics Division, 101:819-838 (December 1975).

16. Potter, J. E. "Matrix Equation Solutions." SIAM Journal on Applied Mathematics, 14:496-501 (May 1966).

17. Kaufman, H. and R. L. Sternberg. "Application of the Theory of Systems of Differential Equations to Multiple Nonuniform Transmission Lines." Journal of Mathematics and Physics, 31:244-252 (1952).

18. Bellman, R., et al. "On the Matrix Riccati Equation of Transport Processes." Journal of Mathematical Analysis and Applications, 44: 472-481 (1973).

19. Vasudevan, R. "On a Generalization of the Matrix and Scalar Riccati Equation." Journal of Mathematical Analysis and Applications, 45: 629-638 (March 1974).

20. Mitra, S. K. "Common Solutions to a Pair of Linear Matrix Equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$." Proceedings of the Cambridge Philosophical Society, 74:213-215 (September 1973).

21. Roth, W. E. "The Equations AX - YB = C and AX - XB = C in Matrices." Proceedings of the American Mathematical Society, 3:392-396 (1952).

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

23. Barnett, S. Introduction to Mathematical Control Theory. London: Oxford University Press, 1975

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

25. Porter, B. "Design of Time-Optimal Regulators for Linear Multivariable Discrete-Time Plants." Electronics Letters, 12:196-197 (April 1976).

26. Hille, E. Lectures on Ordinary Differential Equations. Reading, Massachusetts: Addison-Wesley Publishing Company, 1969.

27. Roth, W. E. "On the Matrix Equation $X^2 + AX + XB + C = 0$." Proceedings of the American Mathematical Society, 1:586-589 (October 1950).

28. Jones, J., Jr., J. Leuthauser, R. Gressang. "Solutions of the Algebraic Riccati Matrix Equation." Notices of the American Mathematical Society, 23:A-517 (August 1976).

29. Martensson, K. "On the Matrix Riccati Equation." Information Sciences, 3:17-49 (1971).

30. Denman, E. D. and A. N. Beavers, Jr. "The Matrix Sign Function and Computations in Systems." Applied Mathematics and Computation, 2: 63-94 (1976).

31. Hagander, P. "Numerical Solution of $A^TS + SA + Q = 0$." Information Sciences, 4:35-50 (1972).

32. Lancaster, P. "Explicit Solutions of Linear Matrix Equations." SIAM Review, 12:544-566 (October 1970).

33. Jennings, W. First Course in Numerical Methods. New York: The Mac-Millan Company, 1969.

34. Liusternik, L. A. and V. J. Sobolev. Elements of Functional Analysis. New York: Frederik Ungar Publishing Company, 1961.

35. Tapia, R. A. "The Kantorovich Theorem for Newton's Method." The American Mathematical Monthly, 78:389-392 (April 1971).

Supplemental Bibliography for the Liapunov Equation

Amaldi, U. "Sulle sostituzioni lineari commutabili." Rendiconti Istituto Lombardo (2), 45:433-455 (1912).

Barnett, S. "Simplification of Certain Linear Matrix Equations." IEEE Transactions on Automatic Control, AC-21:115-116 (February 1976).

Barnett, S. and C. Storey. "Solution of Liapunov Matrix Equation." Electronics Letters, 2:466-467 (December 1966).

-----. "The Liapunov Matrix Equation and Schwarz' Form." IEEE Transactions on Automatic Control, AC-12:117-118 (February 1967).

-----. "Remarks on Numerical Solution of the Liapunov Matrix Equation." Electronics Letters, 3:417-418 (September 1967).

-----. "Some Applications of the Liapunov Matrix Equation." Journal of the Instutitute of Mathematics and its Applications, 4:33-42 (1968).

Bartels, R. H. and G. W. Stewart. "Algorithm 432, Solution of the Matrix Equation AX + XB = C." Communications of the Association for Computing Machinery, 15:820-826 (1972).

Beavers, A. N. and E. D. Denman. "A New Solution Method for the Lyapunov Matrix Equation." SIAM Journal on Applied Mathematics, 29:416-421 (November 1975).

Bingulac, S. P. "An Alternate Approach to Expanding PA + A'P = -Q." IEEE Transactions on Automatic Control, AC-15:135-137 (February 1970).

Brauer, F. and J. Nohel. Qualitative Theory of Ordinary Differential Equations. New York: W. A. Benjamin, Inc., 1969.

Caley, A. "On the Quaternion Equation qQ - Qq' = 0." Messages of Mathematics, 14:108-112 (1885).

-----. "On the Matrix Equation qQ - Qq' = 0." Messages of Mathematics, 14:176-178 (1885).

Cecioni, F. "Sull'equazioni fra matrici AX = XB, X^m = A." Accademia dei Lincei Rendiconti (5), 18:566-571 (1909).

-----. "Sull'equazione fra Matrici AX = XA." Annali Università Toscane 14, 2:1-49 (1931).

Chen, C. F. and L. S. Shieh. "A Note on Expanding PA + A^TP = -Q." IEEE Transactions on Automatic Control, AC-13:122-123 (February 1968).

Davison, E. J. and F. T. Man. "The Numerical Solution of A'Q + QA = -C." IEEE Transactions on Automatic Control, AC-13:448-449 (August 1968).

Frobenius, G. "Über lineare Substitutionen und bilineare Formen." Journal der reinen angewandten Mathematik, 84:1-63 (1878).

-----. "Über die mit einer Matrix vertauschbaren Matrizen." Bericht der Preussischen Akademie der Wissenschaft, 3-15 (1910).

Givens, W. Elementary Divisors and Some Properties of the Lyapunov Mapping X → AX + XA*. Illinois: Argonne National Laboratory, 1961.

Hartfiel, D. J. "The Matrix Equation AXB = X." Pacific Journal of Mathematics, 36:659-670 (1971).

Hartwig, R. E. "The Resultant and the Solution of AX = XB." SIAM Journal on Applied Mathematics, 22:538-544 (June 1972).

Hensel, K. "Theorie der Körper von Matrizen." Journal der reinen angewandten Mathematik, 127:116-166 (1904).

Hodges, J. H. "The Matrix Equation AX = B in a Finite Field." American Mathematical Monthly, 63:243-244 (April 1956).

Jamieson, A. "Solution of the Equation AX + XB = C by Inversion of an MxM or NxN Matrix." SIAM Journal on Applied Mathematics, 16:1020-1023 (September 1968).

Jones, J., Jr. "On the Lyapunov Stability Criteria." SIAM Journal on Applied Mathematics, 13:941-945 (December 1965).

-----. "Explicit Solutions of the Matrix Equation AX - XB = C." Rendiconti del Circolo Matematico di Palermo, 23:245-257 (1974).

Kreisselmeier, G. "A Solution of the Bilinear Matrix Equation AY + YB = -Q." SIAM Journal on Applied Mathematics, 23:334-338 (November 1972).

Kučera, V. "The Matrix Equation AX + XB = C." SIAM Journal on Applied Mathematics, 26:15-25 (January 1974).

Landsberg, G. "Über Fundamentalsysteme und bilineare Formen." Journal der reinen angewandten Mathematik, 116:331-349 (1895).

Lu, C. S. "Solution of the Matrix Equation AX + XB = C." Electronics Letters, 7:185-186 (April 1971).

Luenberger, D. G. "Invertible Solutions to the Operator Equation TA - BT = C." Proceedings of the American Mathematical Society, 16:1226-1229 (1965).

Maurer. Zur Theorie der linearen Substitutionen. Dissertation. Strassburg, 1887.

Porter, A. D. "Solvability of the Matric Equation AX = B." Linear Algebra and its Applications, 13:177-184 (1976).

Rosenblum, M. "On the Operator Equation BX - XA = Q." Duke Mathematical Journal, 23:263-270 (1956).

-----. "The Operator Equation BX - XA = Q with Self-Adjoint A and B." Proceedings of the American Mathematical Society, 20:115-120 (1969).

Roth, W. E. "On the Equation P(A,X) = 0 in Matrices." Transactions of the American Mathematical Society, 35:689-708 (1933).

Rothschild, D. and A. Jamieson. "Comparison of Four Numerical Algorithms for Solving the Liapunov Matrix Equation." International Journal of Control, 11:181-198 (1970).

Rutherford, D. E. "On the Solution of the Matrix Equation AX + XB = C." Akademie van Wetenschappen, Amsterdam, Proceedings, 35:54-59 (1932).

-----. "On the Rational Solution of the Matrix Equation SX = XT." Akademie van Wetenschappen, Amsterdam, Proceedings, 36:432-442 (1933).

Smith, R. A. "Bounds for Quadratic Lyapunov Functions." Journal of Mathematical Analysis and Applications, 12:425-435 (1965).

-----. "Matrix Equations XA + BX = C." SIAM Journal on Applied Mathematics, 16:198-201 (January 1968).

Sylvester, J. J. "Sur l'équation en matrices px = xq." Comptes Rendus Academie Scientifique, Paris, 99:67-71, 115-116 (1884).

Trampus, A. "A Canonical Basis for the Matrix Transformation X → AXB." Journal of Mathematical Analysis and Applications, 14:153-160 (1966).

-----. "A Canonical Basis for the Matrix Transformation X → AX + XB." Journal of Mathematical Analysis and Applications, 14:242-252 (1966).

Voss, A. "Über die mit einer bilinearen Form vertauschbaren Formen." München Bericht, 19:283-300 (1889).

-----. "Synnetrische und alternierende Lösungen der Gleichung SX = XS'." München Bericht, 26:273-281 (1896).

Wedderburn, J. H. "Note on the Linear Matrix Equation." Proceedings of the Edinburgh Mathematical Society, 22:49-53 (1904).

-----. "On Matrices Whose Coefficients are Functions of a Single Variable." Transactions of the American Mathematical Society, 16:328-332 (1915).

Weitzenböck, R. "Über die Matrixgleichung AX + XB = C." Akademie van Wetenshappen, Amsterdam, Proceedings, 35:60-61 (1932).

Supplemental Bibliography for the Matrix Riccati Equation

Aoki, M. "On the Subspaces Associated with Partial Reconstruction of State Vectors, the Structure Algorithm, and the Predictable Directions of Riccati Equations." IEEE Transactions on Automatic Control, AC-18:399-400 (August 1973).

Anderson, J. and J. Parker. "Choices for A in the Matrix Equation T = AB - BA." Linear and Multilinear Algebra, 2:203-209 (1974).

Beavers, A. N. and E. D. Denman. "A New Solution Method of Quadratic Matrix Equations." Mathematical Biosciences, 20:135-143 (June 1974).

-----. "Asymptotic Solutions to the Matrix Riccati Equation." Mathematical Biosciences, 20:339-344 (August 1974).

-----. "A New Similarity Transformation Method for Eigenvalues and Eigenvectors." Mathematical Biosciences, 21:143-169 (October 1974).

-----. "The Matrix Sign Function and Computations in Systems." Applied Mathematics and Computation, 2:63-94 (1976).

Bingulac, S. P., et al. On an Iterative Solution of Time-Invariant Riccati Equation. Joint Automatic Control Conference of the American Automatic Control Council Paper 3-C6. New York: The Institute of Electrical and Electronics Engineers, Inc., 1971.

Caines, P. E. and D. Q. Mayne. "On the Discrete Time Matrix Riccati Equation of Optimal Control." International Journal of Control, 12:785-794 (1970).

Coles, W. J. "Linear and Riccati Systems." Duke Mathematical Journal, 22: 333-338 (1955).

Coppel, W. A. "Matrix Quadratic Equations." Bulletin of the Australian Mathematical Society, 10:377-401 (1974).

Davison, E. J. and M. C. Maki. "The Numerical Solution of the Matrix Riccati Differential Equation." IEEE Transactions on Automatic Control, AC-18: (February 1973).

Denman, E. D. "An Additional Algorithm for a System of Coupled Algebraic Matrix Riccati Equations." IEEE Transactions on Computers, C-25:91-93 (January 1976).

Hager, W. W. and L. L. Horowitz. "Convergence and Stability Properties of the Discrete Riccati Operator Equation and the Associated Optimal Control and Filtering Problems." SIAM Journal on Control and Optimization, 14:295-312 (February 1976).

Jones, E. L. "A Reformulation of the Algebraic Riccati Equation Problem." IEEE Transactions on Automatic Control, AC-21:113-114 (February 1976).

Kalman, R. E. "When Is a Linear Control System Optimal." ASME Journal of Basic Engineering, 86:51-60 (March 1964).

Kalman, R. E. and R. S. Bucy. "New Results in Linear Filtering and Prediction." ASME Journal of Basic Engineering, 83:95-108 (March 1961).

Kleinman, D. L. "On an Iterative Technique for Riccati Equation Computations." IEEE Transactions on Automatic Control, AC-13:114-115 (February 1968).

-----. "Stabilizing a Discrete, Constant, Linear System with Application to Iterative Methods for Solving the Riccati Equation." IEEE Transactions on Automatic Control, AC-19:252-254 (June 1974).

Lainiotis, P. G. "Discrete Riccati Equation Solutions: Partitioned Algorithms." IEEE Transactions on Automatic Control, AC-20:555-556 (August 1975).

Levin, J. J. "On the Matrix Riccati Equation." Proceedings of the American Mathematical Society, 10:519-524 (August 1959).

Levine, W. S. and M. Athans. "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems." IEEE Transactions on Automatic Control, AC-15:44-48 (February 1970).

Martensson, K. Linear Quadratic Control Package Part I, The Continuous Problem. Research Report 6802. Lund, Sweden: Lund Institute of Technology, Division of Automatic Control, April 1968.

O'Donnell, J. Asymptotic Solution of the Matrix Riccati Equation and Optimal Control. Proceedings of the Fourth Annual Allerton Conference on Circuit and System Theory. Urbana, Illinois: October 1966.

Payne, H. J. and L. M. Silverman. "On the Discrete Time Algebraic Riccati Equation." IEEE Transactions on Automatic Control, AC-18:226-234 (June 1973).

Reid, W. T. "A Matrix Differential Equation of the Riccati Type." American Journal of Mathematics, 68:237-246 (1946).

Roberts, J. D. Linear Model Reduction and Solution of Algebraic Riccati Equations by Use of the Sign Function. Report CUED/B-Control/TR13. Cambridge, England: University of Cambridge, 1971.

Sandell, N. R. "On Newton's Method for Riccati Equation Solution." IEEE Transactions on Automatic Control, AC-19:254-255 (June 1974).

Shubert, H. A. "An Analytic Solution for an Algebraic Matrix Riccati Equation." IEEE Transactions on Automatic Control, AC-19:255-256 (June 1974).

Slama, A. and V. Gourishankar. "A Computational Algorithm for Solving a System of Coupled Algebraic Matrix Riccati Equations." IEEE Transactions on Computers, C-23:100-102 (January 1974).

Vaughn, D. R. "A Negative Exponential Solution for Matrix Riccati Equation." IEEE Transactions on Automatic Control, AC-14:72-75 (February 1969).

-----. "A Nonrecursive Algebraic Solution for the Discrete Riccati Equation." IEEE Transactions on Automatic Control, AC-15:597-599 (October 1970).

Whyburn, W. M. "Matrix Differential Equations." American Journal of Mathematics, 56:587-592 (1934).

Willems, J. C. "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation." IEEE Transactions on Automatic Control, AC-16:621-634 (December 1971).

-----. "On the Existence of a Nonpositive Solution to the Riccati Equation." IEEE Transactions on Automatic Control, AC-19:592-593 (October 1974).

Wonham, W. M. On Matrix Quadratic Equations and Matrix Riccati Equations. Technical Report 67-5. Providence, Rhode Island: Division of Applied Mathematics, Brown University, 1967.

-----. "On a Matrix Riccati Equation of Stochastic Control." SIAM Journal of Control, 6:681-697 (November 1968).

Ýakel, R. A. and P. V. Kokotović. "A Boundary Layer Method for the Matrix Riccati Equation." IEEE Transactions on Automatic Control, AC-18: 17-24 (February 1973).

VITA

Captain James L. Leuthauser was born in Des Moines, Iowa, in 1947. After graduating from East High School there, he attended the USAFA from which he received a B. S. degree in Astronautical Engineering in 1969. Subsequent assignments included pilot training at Laredo AFB, Texas, F-4 training, and a tour flying F-4s with the 8th TFW in SEA. He was an F-4 instructor pilot and an F-15 academic instructor at Luke AFB from 1972 until his assignment to the AFIT in May 1975.

Permanent address: 2200 Capitol Avenue
Des Moines, Iowa 50317

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER GA/MA/76D-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* SOLUTIONS OF CERTAIN ALGEBRAIC MATRIX EQUATIONS WITH APPLICATIONS TO ENGINEERING SCIENCES		5. TYPE OF REPORT & PERIOD COVERED MS Thesis 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James L. Leuthauser Captain		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT-EN) Wright-Patterson AFB, Ohio 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Flight Control Division (AFFDL/FG) Air Force Flight Dynamics Laboratory Air Force Wright Aeronautical Laboratories Wright-Patterson AFB, Ohio 45433		12. REPORT DATE December 1976 13. NUMBER OF PAGES 103
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

Approved for public release; IAW AFR 190-17

JERRAL F. GUESS, Captain, USAF
Director of Information

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Matrices (Mathematics)	Linear Algebraic Equations	Quadratic Equations
Matrix Equations	Nonlinear Algebraic Equations	Liapunov Functions
Matrix Factorization	Control Systems	Linear Systems
Linear Equations	Liapunov Equations	Liapunov Stability
Optimal Control Theory	Riccati Equations	

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Solutions of certain matrix equations are presented in this thesis. Results of Mitra are used to extend current methods of obtaining a general form of the solution of the Liapunov equation. Solutions of the general linear matrix equation, including important special cases, are obtained under rather general conditions on the matrices. Necessary and sufficient conditions are established for the existence of a solution of AX = C where the elements of the matrices belong to the polynomial domain of the field of real numbers. Representations

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

Block 19.

Linear Differential Equations
Nonlinear Differential Equations

Block 20.

of solutions to the algebraic Riccati equation are given along with original techniques for computation of solutions and conditions for uniqueness. Sufficient conditions for the existence of the generalized nonlinear algebraic and differential Riccati equations are established. The classical Newton-Raphson method of finding zeros of functions of a single variable is extended to the case of finding zeros of functions of a matrix variable, including determining the inverse of a matrix. An extensive bibliography includes a comprehensive listing of articles which provide historical and developmental background on the matrix equations considered in this thesis.