ON THE MATRIX POLYNOMIAL, LAMBDA-MATRIX AND BLOCK EIGENVALUE PROBLEMS[†]

by

J. E. Dennis, Jr.[*]

J. F. Traub[**]

R. P. Weber[***]

Technical Report 71-109

[*] Department of Computer Science
Cornell University
Ithaca, New York 14850

Supported in part by National Science Foundation Grant GJ-27528

[**] Department of Computer Science
Carnegie - Mellon University
Pittsburgh, Pennsylvania 15213

Supported in part by Office of Naval Research under Contract N 00014-67-A-0314-0010, NR 044-422

[***] Bell Telephone Laboratories, Inc.
Whippany, New Jersey 07981

[†] Also available as a Technical Report from Carnegie-Mellon University Computer Science Department.

ABSTRACT

A matrix S is a solvent of the matrix polynomial

$$M(X) \equiv X^m + A_1X^{m-1} + \cdots + A_m,$$

if $M(S) = \underline{\underline{0}}$, where A_i, X and S are square matrices. We present some new mathematical results for matrix polynomials, as well as a globally convergent algorithm for calculating such solvents.

In the theoretical part of this paper, existence theorems for solvents, a generalized division, interpolation, a block Vandermonde, and a generalized Lagrangian basis are studied.

Algorithms are presented which generalize Traub's scalar polynomial methods, Bernoulli's method, and eigenvector powering.

The related lambda-matrix problem, that of finding a scalar λ such that

$$I\lambda^m + A_1\lambda^{m-1} + \cdots + A_m$$

is singular, is examined along with the matrix polynomial problem.

The matrix polynomial problem can be cast into a block eigenvalue formulation as follows. Given a matrix A of order mn, find a matrix X of order n, such that $AV = VX$, where V is a matrix of full rank. Some of the implications of this new block eigenvalue formulation are considered.

TABLE OF CONTENTS

CHAPTER I

Introduction

In this chapter we state the problem, give some of the definitions, present the major results of the paper, and outline the entire dissertation.

1.1 <u>Preliminaries</u>. Algorithms for the solution of the scalar polynomial problem, $x^m + a_1x^{m-1} + \cdots + a_m = 0$, have become extremely efficient. See Traub [20,21] and Jenkins and Traub [7,8]. A generalization of the scalar polynomial is given by the following.

Definition 1.1 <u>Given n by n matrices $A_0, A_1, \ldots, A_m$, a matrix polynomial $M(X)$ is the matrix function</u>

$$M(X) \equiv A_0X^m + A_1X^{m-1} + \cdots + A_m \tag{1.1}$$

<u>in the n by n matrix variable X.</u>

If A_0 is nonsingular, then the <u>monic</u> matrix polynomial is

$$\bar{M}(X) \equiv A_0^{-1}M(X). \tag{1.2}$$

Two generalizations of the roots of a scalar polynomial are to be examined. The first one, the major emphasis of this work, is classical. Little is known, however, about existence and calculation of such roots of matrix polynomials.

Definition 1.2 *A matrix S is a solvent of the matrix polynomial M(X) if*

$$M(S) = \underline{0}. \qquad (1.3)$$

Definition 1.3 *A matrix W is a weak solvent of the matrix polynomial M(X) if*

$$\det M(W) = 0. \qquad (1.4)$$

A special case of the weak solvent problem is the important lambda-matrix problem. Restricting the class of weak solvents to scalar matrices, λI, and using the notation $M(\lambda) \equiv M(\lambda I)$, the *lambda-matrix* problem is that of finding a scalar λ such that

$$M(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_m \qquad (1.5)$$

is singular. Such a scalar is called a *latent root* of $M(\lambda)$ and vectors b and r are right and left *latent vectors*, respectively, if, for a latent root ρ, $M(\rho)b = \underline{0}$ and $r^T M(\rho) = \underline{0}^T$. $M(\lambda)$ in equation (1.5) is an n by n matrix whose elements are scalar polynomials in λ. See Lancaster [13], Gantmacher [2], MacDuffee [15], and Peters and Wilkinson [17] for a complete discussion of lambda-matrices. A description of some of the present methods of solving the lambda-matrix problem is found in Appendix B.

Only monic matrix polynomials are studied in the main part of this dissertation. The case of the nonmonic matrix polynomial, and where A_0 is singular, will be considered in Appendix A. If A_0 is nonsingular, the monic matrix polynomial $\overline{M}(X)$ can be obtained by the solution of several linear systems, as was suggested by Peters and Wilkinson [17]. Hence, we consider

$$M(X) \equiv X^m + A_1 X^{m-1} + \cdots + A_m. \tag{1.6}$$

The following are some well-known results that will be frequently used. They may all be found in Lancaster [13].

A corollary of Bézout's theorem states that if S is a solvent of M(X) then

$$M(\lambda) \equiv Q(\lambda)(I\lambda-S), \tag{1.7}$$

where $Q(\lambda)$ is a monic lambda-matrix of degree m-1. Another result is that the lambda-matrix $M(\lambda)$ has mn latent roots, and hence, it follows immediately from (1.7) that the n eigenvalues of a solvent are all latent roots of the lambda-matrix. Furthermore, the n(m-1) latent roots of $Q(\lambda)$ are also latent roots of $M(\lambda)$.

If one is interested in the solution of a lambda-matrix problem, then a solvent will provide n latent roots and can be used for a matrix deflation, which yields the new problem $Q(\lambda)$.

1.2 Main Results of this Paper. The following are the principal results of this work. They will be proved in later chapters.

The Fundamental Theorem of Algebra, that a scalar polynomial has at least one zero, does not hold true for matrix polynomials. There are matrix polynomials which have no solvents (Theorem 2.6).

It is useful to have a concept of a matrix polynomial with a complete set of solvents. This is a generalization of an n^{th} degree scalar polynomial having n roots.

Definition 1.4 A set of m solvents of $M(X)$ is a complete set of solvents, if the set of mn eigenvalues of the m solvents is the same, counting multiplicities, as the set of mn latent roots of $M(\lambda)$.

Thus, in the special case of $M(\lambda)$ having mn distinct latent roots, a complete set of m solvents must have no common eigenvalues and each solvent must have distinct eigenvalues.

We consider a generalization of the scalar Vandermonde matrix.

Definition 1.5 Given n by n matrices $S_1,\cdots,S_m$, the block Vandermonde matrix is

$$V(S_1,\cdots,S_m) \equiv \begin{pmatrix} I & I & \cdots & \cdots & I \\ S_1 & S_2 & \cdots & \cdots & S_m \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ S_1^{m-1} & S_2^{m-1} & \cdots & \cdots & S_m^{m-1} \end{pmatrix} \quad (1.8)$$

It will be shown in Chapter 4 that it is not sufficient that matrices $S_1,\cdots,S_m$ have distinct and disjoint eigenvalues for $V(S_1,\cdots,S_m)$ to be nonsingular.

Existence of a complete set of solvents for the important special case of the lambda-matrix having distinct latent roots is given by the following theorem (Theorem 4.1).

Theorem <u>If $M(\lambda)$ has distinct latent roots, then $M(X)$ has a complete set of solvents, $S_1,\cdots,S_m$, and $V(S_1,\cdots,S_m)$ is nonsingular.</u>

Definition 1.6 <u>A solvent of $M(X)$ is a dominant solvent if the n eigenvalues of this solvent are strictly the n largest latent roots of $M(\lambda)$.</u>

Algorithm 1, presented below and again in Chapter 5, attempts to find a dominant solvent of $M(X)$. It is a generalization of one due to Traub [21] for scalar polynomials. The algorithm has two stages. The first, a generalization of Sebastião e Silva's algorithm (see Householder [4]), generates a sequence of matrix polynomials, all of degree less than m. Then the last two matrix polynomials of the generated sequence are used in a matrix iteration which is to converge to a dominant solvent.

Algorithm 1

(i) *Let* $G_o(X) = I$ *and*

$$G_{n+1}(X) = G_n(X)X - \alpha_1^n M(X), \tag{1.9}$$

for $n = 0,1,\cdots,L-1,$ *where*

$$G_n(X) \equiv \alpha_1^n X^{m-1} + \cdots + \alpha_m^n. \tag{1.10}$$

(ii) *Let* $X_o = \left(\alpha_1^L\right)\left(\alpha_1^{L-1}\right)^{-1}$ *and*

$$X_{i+1} = G_L(X_i)G_{L-1}^{-1}(X_i) \tag{1.11}$$

for $i = 0,1,\cdots.$

Convergence of this algorithm is established for a class of matrix polynomials (Theorem 5.1).

Theorem *If*

(i) $M(X)$ *has a complete set of solvents*, $S_1,\cdots,S_m$,

(ii) S_1 *is a dominant solvent*, *and*,

(iii) $V(S_1,\cdots,S_m)$ *and* $V(S_2,\cdots,S_m)$ *are nonsingular*,

then

(i) $\overline{G}_n(X) \equiv \left(\alpha_1^n\right)^{-1} G_n(X) \to \overline{M}_1(X)$ *as* $n \to \infty$, *where*

<u>$\overline{M}_1(X)$ is the unique monic matrix polynomial of degree m-1 with solvents $S_2,\cdots,S_m$, but not S_1, and</u>

(ii) <u>for L sufficiently large X_1 of equation (1.11) converges to S_1</u>

It will be shown (Corollary 5.2 and Lemma 5.7) that each stage of the algorithm is linearly convergent. Let σ be the absolute value of the ratio of the smallest eigenvalue of S_1 and the largest remaining latent root of $M(\lambda)$. Then the asymptotic error constants of the first and second stage are $c_1\sigma$ and $c_2\sigma^{L-1}$, respectively, where $\sigma < 1$ and L is the number of iterations of the first stage before switching to the second stage. Thus, the second stage, though linearly convergent, can be made arbitrarily fast by increasing the number of iterations of the first stage. In the computational algorithm, we pick an arbitrary L and then examine the second stage. If it is converging too slowly (or diverging), then the first stage is resumed for several steps and the process is continued. Thus, given that the three hypotheses of the above theorem are satisfied, this process, in exact arithmetic, is guaranteed to yield a solvent of the matrix polynomial.

If a dominant solvent does not exist, then the algorithm will not yield a solvent. In addition to the results in the above theorem, the first stage yields a dominant latent root, if one exists. Consider the following algorithm which obtains a dominant latent root (Chapter 7).

Definition 1.7 *Given vectors* $v_o, v_1, \cdots, v_m$ *of dimension* n, *a lambda-vector* $g(\lambda)$ *is the vector function*

$$g(\lambda) \equiv v_o\lambda^m + v_1\lambda^{m-1} + \cdots + v_m. \tag{1.12}$$

Algorithm 2 *Let* $g_o(\lambda)$ *be an arbitrary* m-1 *degree lambda-vector. Generate*

$$g_{n+1}(\lambda) = g_n(\lambda)\lambda - M(\lambda)v_1^{(n)}, \tag{1.13}$$

where

$$g_n(\lambda) \equiv v_1^{(n)}\lambda^{m-1} + \cdots + v_m^{(n)}. \tag{1.14}$$

This is another generalization of Traub's scalar polynomial algorithm. For a vector v, denote by max v the first element of v which has the maximum absolute value. Note that max v is not a norm. Then a convergence theorem for the algorithm is as follows (Theorem 7.1).

Theorem *If*

(i) $M(\lambda)$ *has distinct latent roots*, $\rho_1, \cdots, \rho_{mn}$,

(ii) $|\rho_1| > |\rho_i|$ *for* $i \neq 1$, *and*

(iii) $r_1^T g_o(\rho_1) \neq 0$, *where* $r_1^T M(\rho_1) = \underline{0}^T$,

then

(i) $\bar{g}_n(\lambda) \equiv \dfrac{g_n(\lambda)}{\max v_1^{(n)}} \to \dfrac{M(\lambda)}{\lambda - \rho_1} b_1$, *where* $M(\rho_1)b_1 = \underline{0}$

and

$$\text{(ii)} \quad \frac{v_1^{(n+1)} - \rho_1 v_1^{(n)}}{\max v_1^{(n)}} \to \underline{0}. \tag{1.15}$$

The transpose of any column of equation (1.9) with $X = \lambda I$, is precisely equation (1.13), with $M^T(\lambda)$ replacing $M(\lambda)$. Since the latent roots of $M^T(\lambda)$ are the same as those of $M(\lambda)$, a dominant latent root of $M(\lambda)$ can be obtained from equation (1.15) by Algorithm 1, the matrix polynomial solvent algorithm. This can be done regardless of whether a dominant solvent, or any solvent at all, exists.

1.3 Outline of the Remainder of the Paper. This paper contains three intertwined yet distinct subjects. They are

(i) new theoretical results on matrix polynomials,

(ii) algorithms for solvents and latent roots, and

(iii) a new block eigenvalue problem.

Chapter 2 considers the basic properties of solvents. The existence of solvents and factorization of lambda-matrices are considered here. A generalization of Bézout's Theorem and the relationship between polynomial coefficients and the elementary symmetric functions are also discussed.

In Chapter 3 we present some of the basic properties of matrix polynomials. Interpolation, representation theorems and fundamental matrix polynomials are presented in this chapter.

Properties of the block Vandermonde matrix are given in Chapter 4.

The second major area of this dissertation concerns itself with algorithms for finding solvents and latent roots. Chapter 5 presents Algorithm 1, the main algorithm of the paper. The method finds solvents and is a generalization of Traub's scalar polynomial methods [21]. A convergence theorem, computational discussion and flow-chart are given here.

A block Bernoulli method is described in Chapter 6. The relation between this method and Algorithm 1 is discussed.

In Chapter 7 we present Algorithm 2, which finds a dominant latent root. The key result is given - the computations of Algorithm 2 are done by Algorithm 1. A vector Bernoulli method is also described.

The third area of this work is a new block eigenvalue problem. It is that of finding a matrix X of order n such that for given matrix A of order mn, the equation $AV = VX$ is satisfied for a matrix V of full rank. Chapter 8 deals with this problem. It is shown that when A is the block companion matrix, this problem is a generalization of the matrix polynomial solvent problem. A general theory of block eigenvalues as well as two algorithms based on eigenvector powering are offered.

Chapter 9 describes numerical testing of Algorithms 1 and 2.

CHAPTER 2

Solvents

In this chapter we study some of the properties of solvents. Section 2.1 considers a division of matrix polynomials which results in a new derivation and generalization of Bézout's theorem. Section 2.2 examines the block companion matrix. Principal vectors of solvents are considered in Section 2.3. The existence of solvents and factorization of lambda-matrices are both dealt with in Section 2.4.

2.1 Generalized Division. The class of matrix polynomials is not closed under multiplication or division. Consider the product of $N(X) \equiv X + N$ and $L(X) \equiv X + L$. We get $N(X)L(X) = (X+N)(X+L) = X^2 + NX + XL + NL$ which is not of the general form of a matrix polynomial; $X^2 + A_1X + A_2$. A new operation will be defined for matrix polynomials which will reduce to division in the scalar case; $n = 1$.

Theorem 2.1 Let $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$ and $W(X) = X^p + B_1X^{p-1} + \cdots + B_p$, with $m \geq p$. Then there exists a unique, monic matrix polynomial $F(X)$ of degree $m-p$ and a unique matrix polynomial $L(X)$ of degree $p-1$ such that

$$M(X) \equiv F(X)X^p + B_1F(X)X^{p-1} + \cdots + B_pF(X) + L(X). \quad (2.1)$$

Proof: Let $F(X) = X^{m-p} + F_1X^{m-p-1} + \cdots + F_{m-p}$ and $L(X) = L_0X^{p-1} + L_1X^{p-2} + \cdots + L_{p-1}$. Equating coefficients of equation (2.1), $F_1, F_2, \cdots, F_{m-p}$ and $L_0, L_1, \cdots, L_{p-1}$ can be successively and uniquely determined from the m equations. #

Equation (2.1) is the matrix polynomial division of $M(X)$ by $W(X)$ with quotient $F(X)$ and remainder $L(X)$.

Definition 2.1 Associated with the matrix polynomial, $M(X) \equiv X^m + A_1X^{m-1} + \cdots + A_m$, is the commuted matrix polynomial

$$\hat{M}(X) \equiv X^m + X^{m-1}A_1 + \cdots + A_m. \tag{2.2}$$

If $\hat{M}(R) = \underline{0}$, then R is a left solvent of $M(X)$.

The matrix S such that $M(S) = \underline{0}$, previously just called a solvent, will be referred to as a right solvent when confusion might occur.

An important association between the remainder, $L(X)$, and the dividend, $M(X)$, in equation (2.1), will now be given. It generalizes the fact that for scalar polynomials the dividend and remainder are equal when evaluated at the roots of the divisor.

Corollary 2.1 If R is a left solvent of $W(X)$, then $\hat{L}(R) = \hat{M}(R)$.

Proof: Let $Q(X) = M(X) - L(X)$. Then, it is easily shown that

$$\hat{Q}(X) \equiv X^{m-p}\hat{W}(X) + X^{m-p-1}\hat{W}(X)F_1 + \cdots + \hat{W}(X)F_{m-p}. \tag{2.3}$$

The result immediately follows since $\hat{Q}(R) = \underline{0}$ for all left solvents of $W(X)$. #

The case where $p = 1$ is very useful in this paper. Here we have $W(X) = X - R$ where R is both a left and right solvent of $W(X)$. Then Theorem 2.1 shows that

$$M(X) \equiv F(X)X - RF(X) + L \qquad (2.4)$$

where L is a constant matrix. Now Corollary 2.1 shows that $L = \hat{M}(R)$, and, thus,

$$M(X) \equiv F(X)X - RF(X) + \hat{M}(R). \qquad (2.5)$$

There is a corresponding theory for $\hat{M}(X)$. In this case, equation (2.1) is replaced by

$$\hat{M}(X) \equiv X^p\hat{H}(X) + X^{p-1}\hat{H}(X)B_1 + \cdots + \hat{H}(X)B_p + \hat{N}(X) \quad (2.6)$$

and Corollary 2.1 becomes the following.

Corollary 2.2 _If_ S _is a right solvent of_ $W(X)$, _then_ $N(S) = M(S)$.

We again consider the case of $p = 1$. Let $W(X) = X - S$. Then equation (2.5) becomes

$$\hat{M}(X) \equiv X\hat{H}(X) - \hat{H}(X)S + M(S). \qquad (2.7)$$

Restricting X to a scalar matrix λI, and noting that

$M(\lambda) \equiv \hat{M}(\lambda)$, we get Bézout's Theorem (see Gantmacher [2, vol. I, p. 81]) from equations (2.5) and (2.7):

$$M(\lambda) \equiv (I\lambda-R)F(\lambda) + \hat{M}(R) \equiv H(\lambda)(I\lambda-S) + M(S) \qquad (2.8)$$

for any matrices R and S. If in addition R and S are left and right solvents, respectively, of $M(X)$, then

$$M(X) \equiv F(X)X - RF(X), \qquad (2.9)$$

$$\hat{M}(X) \equiv X\hat{H}(X) - \hat{H}(X)S \qquad (2.10)$$

and

$$M(\lambda) \equiv (I\lambda-R)F(\lambda) \equiv H(\lambda)(I\lambda-S). \qquad (2.11)$$

Hence, Corollaries 2.1 and 2.2 are generalizations of Bézout's Theorem.

The use of block matrices is fundamental in this work. For notational purposes it is useful to have a concept of the transpose of a block matrix without transposing the blocks.

Definition 2.2 _Let_ A _be a matrix with block structure_ (B_{ij}) _with_ B_{ij} _matrices of order_ n. _The block transpose of dimension_ n _of_ A, _denoted_ $A^{B(n)}$, _is the matrix with block structure_ (B_{ji}).

The order of the block transpose will generally be dropped when it is clear. Note that, in general, $A^{B(n)} \neq A^T$, except when $n = 1$.

A scalar polynomial exactly divides another scalar polynomial, if all the roots of the divisor are roots of the dividend. A generalization of the scalar polynomial result is given next. The notation is that of Theorem 2.1.

Corollary 2.3 <u>If</u> $W(X)$ <u>has</u> p <u>left solvents</u>, $R_1,\cdots,R_p$ <u>which are also left solvents of</u> $M(X)$, <u>and if</u> $V^B(R_1,\cdots,R_p)$ <u>is nonsingular, then the remainder</u> $L(X) \equiv \underline{0}$.

Proof: Corollary 2.1 shows that $\hat{L}(R_i) = \underline{0}$ for $i = 1,\cdots,p$. Since $V^B(R_1,\cdots,R_p)$ is nonsingular, and since

$$\begin{pmatrix} I & R_1 & \cdots & R_1^{p-1} \\ I & R_2 & \cdots & R_2^{p-1} \\ \vdots & \vdots & & \vdots \\ I & R_p & \cdots & R_p^{p-1} \end{pmatrix} \begin{pmatrix} L_{p-1} \\ L_{p-2} \\ \vdots \\ L_o \end{pmatrix} = \begin{pmatrix} \hat{L}(R_1) \\ \hat{L}(R_2) \\ \vdots \\ \hat{L}(R_p) \end{pmatrix} = \underline{0}$$

it follows that $L(X) \equiv \underline{0}$. Thus,

$$M(X) \equiv F(X)X^p + B_1F(X)X^{p-1} + \cdots + B_pF(X). \quad \# \quad (2.12)$$

From equation (2.11) it follows that the eigenvalues of any solvent (left or right) of $M(X)$ are latent roots of $M(\lambda)$. These equations allow us to think of right (left) solvents of $M(X)$ as right (left) factors of $M(\lambda)$.

In the scalar polynomial case, due to commutivity, right and left factors are equivalent. Relations between left and right solvents can now be given.

Corollary 2.4 *If S_j and R_i are right and left solvents of $M(X)$, respectively, and S_j and R_i have no common eigenvalues, then $F_i(S_j) = \underset{\sim}{0}$, where $F_i(X)$ is $F(X)$ defined by equation (2.9) with $R = R_i$.*

Proof: Equation (2.9) shows that

$$F_i(S_j)S_j - R_iF_i(S_j) = \underset{\sim}{0}. \tag{2.13}$$

Since S_j and R_i have no common eigenvalues, $F_i(S_j) = \underset{\sim}{0}$ uniquely. This follows, since the solution of $AX = XB$ has the unique solution $X = \underset{\sim}{0}$, if and only if A and B have no common eigenvalues. See Gantmacher [2, p. 215]. #

Given a left solvent R_i of $M(X)$, Theorem 2.1 shows that $F_i(X)$ exists uniquely. If S is a right solvent of $M(X)$ and if $F_i(S)$ is nonsingular (S is not a weak solvent of $F_i(X)$), then equation (2.13) shows that

$$R_i = F_i(S)SF_i^{-1}(S). \tag{2.14}$$

This gives an association between left and right solvents.

2.2 *Block Companion Matrix.* A useful tool in the study of scalar polynomials is the companion matrix. The eigenvalues of a companion matrix are the roots of its associated polynomial. See Wilkinson [22, p. 12]. A generalization of this is given below. Definition 2.3, Theorem 2.2 and Corollary 2.5 can be found in Lancaster [13].

Definition 2.3 <u>Given a matrix polynomial</u>

$$M(X) \equiv X^m + A_1X^{m-1} + \cdots + A_m,$$

<u>the block companion matrix associated with it is</u>

$$C = \begin{pmatrix} 0 & \cdots & 0 & -A_m \\ I & & & -A_{m-1} \\ & \ddots & & \vdots \\ & & I & -A_1 \end{pmatrix} \tag{2.15}$$

It is well known that the eigenvalues of the block companion matrix are latent roots of the associated lambda-matrix. See Wilkinson [22, p. 12]. Simple algebraic manipulation yields this result.

Theorem 2.2 $\mathrm{Det}(C-\lambda I) \equiv (-1)^{mn}\det\left(I\lambda^m+A_1\lambda^{m-1}+\cdots+A_m\right)$.

Since C is an mn by mn matrix, we immediately obtain the following.

Corollary 2.5 $M(\lambda)$ <u>has exactly</u> mn <u>finite latent roots</u>.

The form of the block companion matrix could have been chosen differently. Theorem 2.2 also holds for the block transpose of the companion matrix:

$$C^B = \begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix}. \qquad (2.16)$$

The algorithms given in this paper are based on eigenvector powering schemes. It will be useful to know the eigenvectors of the block companion matrix and its block transpose. The results are a direct generalization of the scalar case.

Theorem 2.7 <u>If ρ_1 is a latent root of $M(\lambda)$ and b_1 and r_1 are right and left latent vectors, then ρ_1 is an eigenvalue of</u> C <u>and of</u> C^B <u>and</u>

(i) $\begin{pmatrix} b_1 \\ \rho_1 b_1 \\ \vdots \\ \rho_1^{m-1} b_1 \end{pmatrix}$ <u>is the right eigenvector of</u> C^B,

(ii) $\begin{pmatrix} r_1 \\ \rho_1 r_1 \\ \vdots \\ \rho_1^{m-1} r_1 \end{pmatrix}$ <u>is the left eigenvector of</u> C, <u>and</u>

(iii) $\begin{pmatrix} b_i^{(m-1)} \\ \vdots \\ b_i^{(1)} \\ b_i \end{pmatrix}$ <u>is the right eigenvector of</u> C, <u>where</u>

$$\frac{M(\lambda)b_i}{\lambda - \rho_i} \equiv b_i\lambda^{m-1} + b_i^{(1)}\lambda^{m-2} + \cdots + b_i^{(m-1)}. \qquad (2.17)$$

Proof: Parts (i) and (ii) are easily verified by substitutions into the appropriate eigenvalue problem. For part (iii), consider

$$\begin{pmatrix} 0 & \cdots & 0 & -A_m \\ I & & & -A_{m-1} \\ & \ddots & & \vdots \\ & & I & -A_1 \end{pmatrix} \begin{pmatrix} d_i^{(m-1)} \\ \vdots \\ d_i^{(1)} \\ d_i^{(0)} \end{pmatrix} = \rho_i \begin{pmatrix} d_i^{(m-1)} \\ \vdots \\ d_i^{(1)} \\ d_i^{(0)} \end{pmatrix}. \qquad (2.18)$$

Multiply out; multiply the j^{th} component equation by λ^{j-1}; and add. The result is

$$H_i(\lambda)\lambda - M(\lambda)d_i^{(0)} \equiv \rho_i H_i(\lambda), \qquad (2.19)$$

where

$$H_i(\lambda) \equiv d_i^{(m-1)} + d_i^{(m-2)}\lambda + \cdots + d_i^{(0)}\lambda^{m-1}. \qquad (2.20)$$

Equation (2.19) at $\lambda = \rho_i$ shows that $M(\rho_i)d_i^{(0)} = \underline{0}$ and, hence, $d_i^{(0)}$ is a right latent vector. Manipulating equation (2.19), the result equation (2.17) with $d_i^{(0)} = b_i$ and $d_i^{(j)} = b_i^{(j)}$ for $j = 1,\cdots,m-1$, follows. #

2.3 Structure of Solvents. The eigenvectors and principal vectors of a solvent will now be considered. From equation (2.11) it follows that the eigenvectors of a left (right) solvent are left (right) latent vectors of the lambda-matrix. Lancaster [13, p. 50] gives the characterization of a solvent that has only elementary divisors.

Theorem 2.4 *If $M(\lambda)$ has n linearly independent right latent vectors, $b_1,\cdots,b_n$, corresponding to latent roots $\rho_1,\cdots,\rho_n$, then $Q\Lambda Q^{-1}$ is a right solvent, where $Q = [b_1,\ldots,b_n]$ and $\Lambda = \mathrm{diag}(\rho_1,\ldots,\rho_n)$.*

Proof: From $M(Q\Lambda Q^{-1}) = \left(Q\Lambda^m + A_1 Q\Lambda^{m-1} + \cdots + A_m Q\right)Q^{-1}$ the result follows, since $Q\Lambda^m + A_1 Q\Lambda^{m-1} + \cdots + A_m Q$ is just $M(\rho_i)b_i = \underline{0}$ for $i = 1,\cdots,n$. #

It follows from the above proof that if a solvent is diagonalizable, then it must be the form $Q\Lambda Q^{-1}$, as in the above theorem.

Corollary 2.6 *If $M(\lambda)$ has mn distinct latent roots, and the set of right latent vectors satisfy the Haar condition (that every set of n of them are linearly independent), then there are exactly $\binom{mn}{n}$ different right solvents.*

Consider next the case of a solvent which is not diagonalizable. In a manner similar to Roth [18], we consider the principal vectors of a solvent.

Definition 2.4 <u>The j^{th} principal latent vectors of $M(\lambda)$ with respect to the latent root ρ is P_j, which satisfies</u>

$$\frac{1}{(j-1)!} M^{(j-1)}(\rho)P_1 + \frac{1}{(j-2)!} M^{(j-2)}(\rho)P_2 + \cdots + M(\rho)P_j = \underline{0}, \tag{2.21}$$

<u>where</u>

$$M^{(k)}(\lambda) = \frac{d^k}{d\lambda^k} M(\lambda).$$

Note that the first principal latent vector is a latent vector.

Theorem 2.5 <u>The principal vectors of a solvent are principal latent vectors of $M(\lambda)$.</u>

Proof: To alleviate notational difficulties, consider the case where $m = 2$ and $n = k = 3$. The Jordan form of the solvent is $J = \begin{pmatrix} \rho & 1 & \\ & \rho & 1 \\ & & \rho \end{pmatrix}$. Let $P = (P_1 P_2 P_3)$ where $S = PJP^{-1}$ is the solvent of $M(X) = X^2 + A_1 X + A_2$. Thus,

$$\underline{Q} = M(S)P = \left[(P_1P_2P_3)J^2 + A_1(P_1P_2P_3)J + A_2(P_1P_2P_3)\right]$$

$$= \Big[\left(I\rho^2+A_1\rho+A_2\right)P_1, (2I\rho+A_1)P_1 + \left(I\rho^2+A_1\rho+A_2\right)P_2, IP_1$$

$$+ (2I\rho+A_1)P_2 + \left(I\rho^2+A_1\rho+A_2\right)P_3\Big]$$

$$= \Big[M(\rho)P_1, M'(\rho)P_1$$

$$+ M(\rho)P_2, \frac{1}{2}M''(\rho)P_1 + M'(\rho)P_2 + M(\rho)P_3\Big].$$

Hence, P_1, P_2 and P_3, the principal vectors of S, satisfy equation (2.21), the definition of principal latent vectors. #

It is the strategy of this paper to solve the lambda-matrix problem by finding solvents and then finding the eigenvalues of those solvents. The calculation of solvents from the solution of the latent root problem has been considered in the literature. The following is a short description of the method.

Since the eigenvalues of a solvent are latent roots of the lambda-matrix, and there are mn latent roots, it follows that there are only a finite number of Jordan forms of potential solvents. Let the latent roots be given and let J be a matrix in Jordan form with n of the latent roots as its eigenvalues. Then, to find a corresponding solvent S, if one exists, a nonsingular matrix P must be found such that $M(PJP^{-1}) = \underline{0}$. Thus, a nonsingular matrix P must be found such that

$$PJ^m + A_1PJ^{m-1} + \cdots + A_mP = \underline{0}. \tag{2.22}$$

This approach, described in MacDuffee [15, p. 95], is of the general form

$$A_1XB_1 + A_2XB_2 + \cdots + A_mXB_m = C. \tag{2.23}$$

Lancaster [11] and Gantmacher [2] have considered the solution of equation (2.23). The problem is difficult numerically.

Algorithm 1 tries to find a solvent directly, rather than by the above route of solving the latent root problem first.

2.4 <u>Existence of Solvents</u>. We now show that the Fundamental Theorem of Algebra does not hold for matrix polynomials.

Theorem 2.6 <u>There exists a matrix polynomial with no solvents.</u>

Proof: Consider

$$M(\lambda) = \begin{pmatrix} \lambda^2-2\lambda+2 & 1 \\ -1 & \lambda^2-2\lambda \end{pmatrix} = I\lambda^2 - 2I\lambda + \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.24}$$

Det $M(\lambda) = \lambda^4 - 4\lambda^3 + 6\lambda^2 - 4\lambda + 1$, which has all four roots at $\lambda = 1$. Thus, the Jordan form of a solvent must either be $J_1 = I$ or $J_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Since $M(I) \neq \underline{0}$, it follows that J_2 is the only feasible Jordan form. $M(1) = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ and, thus, $b = (1,-1)^T$ is the only latent vector, to within a scalar multiple. The second principal vector is such that $M'(1)b + M(1)P_2 = \underline{0}$. Here,
$M'(\lambda) = \begin{pmatrix} 2\lambda-2 & 0 \\ 0 & 2\lambda-2 \end{pmatrix}$ and, hence, $M'(1) = \underline{0}$.
Thus, $P_2 = b$ to within a scalar multiple. Using Theorem 2.5 and the linear dependence of the first two principal latent vectors, it follows that J_2 is not a feasible Jordan form for a solvent of equation (2.24). #

Consider now the special case of a matrix polynomial whose associated lambda-matrix has distinct latent roots. It will be shown that in this case a complete set of solvents always exists. First we need the following fact about block matrices.

Lemma 2.1 <u>If a matrix</u> A <u>is nonsingular, then there exists a permutation of the columns of</u> A <u>to</u> $\tilde{A}$ <u>such that</u> $A = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$, <u>with</u> $\tilde{A}_{11}$ <u>and</u> $\tilde{A}_{22}$ <u>nonsingular</u>.

Proof: Let A and $\tilde{A}_{11}$ be matrices of orders n and k, respectively, with arbitrary $1 \leq k < n$. Assume the lemma is false. Consider evaluating the determinant as follows. For each of the first k rows, pick an element from a different column. Then multiply these elements and the remaining minor. The sum, with appropriate signs, of every possible choice of the k columns, is the determinant of A. The k choices of the columns determine a square matrix. If that matrix is nonsingular, then the minor must be zero, since the lemma was assumed false. Thus, such terms make no contribution to the determinant of A. A particular minor appears several times in the sum. It occurs the number of ways the same k columns can be picked in different orders. Each minor can thus be factored from several terms; the result being the minor times the determinant of the matrix formed by the k columns and the first k rows. Thus, if the matrix formed by the k columns is singular, then there is no contribution from this term in the determinant of A. Therefore, A must be singular, which is a contradiction. #

Once the columns of A are permutated to get $\tilde{A}_{11}$ and $\tilde{A}_{22}$ nonsingular, the process can be continued to similarly divide $\tilde{A}_{22}$ into nonsingular blocks without destroying the nonsingularity of $\tilde{A}_{11}$.

Theorem 2.7 If A, a matrix of order mn, is nonsingular, then there exists a permutation of the columns of A to $\tilde{A} = (B_{ij})$, with B_{ij} a matrix of order n, such that B_{ii} is nonsingular for $i = 1,\cdots,m$.

The important existence theorem is now given.

Theorem 2.8 *If the latent roots of $M(\lambda)$ are distinct, then $M(X)$ has a complete set of solvents.*

Proof: If the latent roots of $M(\lambda)$ are distinct, then the eigenvalues of the block companion matrix are distinct, and, hence, the eigenvectors of the block companion matrix are linearly independent. From Theorem 2.3 the set of vectors $\begin{pmatrix} b_i \\ \rho_i b_i \\ \vdots \\ \rho_i^{m-1} b_i \end{pmatrix}$ for for $i = 1,\cdots,mn$ are eigenvectors of C^B. The matrix whose columns are these mn vectors is nonsingular. Theorem 2.7 shows that there are m disjoint sets of n linearly independent vectors b_i. Using the structure $Q\Lambda Q^{-1}$ of Theorem 2.4, the complete set of solvents can be formed. #

Corollary 2.7 *If $M(\lambda)$ has distinct latent roots, then it can be factored into the product of linear lambda-matrices.*

Proof: Since $M(\lambda)$ has distinct latent roots, there exists a right solvent S and $M(\lambda) = Q(\lambda)(I\lambda-S)$. $Q(\lambda)$ has the remaining latent roots of $M(\lambda)$ as its latent roots. It follows then, that the latent roots of $Q(\lambda)$ are distinct. Thus, the process can be continued until the last quotient is linear. #

The process described in the above proof considers solvents of the sequence of lambda-matrices formed by the division $M(\lambda) = Q(\lambda)(I\lambda-S)$.

Definition 2.5 <u>A sequence of matrices</u> $C_1,\cdots,C_m$ <u>form a chain of solvents of</u> $M(X)$ <u>if</u> C_i <u>is a right solvent of</u> $Q_i(X)$, <u>where</u> $Q_m(X) \equiv M(X)$ <u>and</u>

$$Q_i(\lambda) \equiv Q_{i-1}(\lambda)(I\lambda-C_i). \tag{2.25}$$

It should be noted that, in general, only C_m is a right solvent of $M(X)$. Furthermore, C_1 is a left solvent of $M(X)$. An equivalent definition of a chain of solvents could be defined with C_i, a left solvent of $T_i(X)$, and

$$T_i(\lambda) \equiv (I\lambda-C_{m-i+1})T_{i-1}(\lambda). \tag{2.26}$$

Corollary 2.8 <u>If</u> $M(\lambda)$ <u>has distinct latent roots, then</u> $M(X)$ <u>has a chain of solvents.</u>

Given C_i and $Q_i(\lambda)$, $Q_{i-1}(\lambda)$ of equation (2.25) can be found by a generalized Horner division scheme. In the numerical solution of the lambda-matrix problem, the strategy considered here will be to find a chain of solvents using the matrix polynomial solvent algorithm and Horner division.

If $C_1,\cdots,C_m$ form a chain of solvents of $M(X)$, then

$$M(\lambda) \equiv I\lambda^m + A_1\lambda^{m-1} + \cdots + A_m \equiv (I\lambda-C_1)(I\lambda-C_2)\cdots(I\lambda-C_m). \tag{2.27}$$

This leads to a generalization of the classical result for scalar polynomials which relates coefficients to elementary symmetric functions. By equating coefficients of equation (2.27) one gets the following theorem.

Theorem 2.9 *If $C_1, \cdots, C_m$ form a chain of solvents for $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$, then*

$$\begin{aligned} A_1 &= -(C_1+C_2+\cdots+C_m) \\ A_2 &= (C_1C_2+C_1C_3+\cdots+C_{m-1}C_m) \\ &\vdots \\ A_m &= (-1)^m C_1C_2\cdots C_m. \end{aligned} \tag{2.28}$$

CHAPTER 3

Properties of Matrix Polynomials

Some of the basic properties of matrix polynomials are considered in this chapter. Section 3.1 concerns itself with matrix polynomial interpolation. A generalization of the fundamental scalar polynomials is given. Representation theorems for matrix polynomials, lambda-matrices, and lambda-vectors are presented in Section 3.2. Section 3.3 studies the fundamental matrix polynomials.

3.1 Interpolation. Given scalars $s_1,\cdots,s_m$, the fundamental polynomials $m_i(x) \equiv \dfrac{p(x)}{(x-s_i)p'(s_i)}$, where $p(x) \equiv \prod_{i=1}^{m} (x-s_i)$, are of great importance in interpolation theory. Their usefulness comes from the fact that $m_i(s_j) = \delta_{ij}$. We will now generalize this for our matrix problem.

Definition 3.1. Given a set of matrices, $S_1,\cdots,S_m$, the fundamental matrix polynomials are a set of m-1 degree matrix polynomials, $M_1(X),\cdots,M_m(X)$, such that $M_i(S_j) = \delta_{ij}I$.

Sufficient conditions, on the set of matrices $S_1,\cdots,S_m$, for a set of fundamental matrix polynomials to exist uniquely will be given in Theorem 3.2. First, however, we need the following results.

Theorem 3.1 Given m pairs of matrices, (X_i,Y_i) $i = 1,\cdots,m$, then there exists a unique matrix polynomial

$P(X) = A_1X^{m-1} + A_2X^{m-2} + \cdots + A_m$, *such that* $P(X_i) = Y_i$ *for* $i = 1,\cdots,m$, *if and only if* $V(X_1,\cdots,X_m)$ *is nonsingular.*

Proof: $P(X_i) = Y_i$ for $i = 1,\cdots,m$ is equivalent to

$$(A_m, A_{m-1}, \cdots, A_1)\begin{pmatrix} I & \cdots & I \\ X_1 & \cdots & X_m \\ \vdots & & \vdots \\ X_1^{m-1} & \cdots & X_m^{m-1} \end{pmatrix} = (Y_1, Y_2, \cdots, Y_m). \quad \#$$

Corollary 3.1 *Given m pairs of matrices,* (X_i, Y_i), $i = 1,\cdots,m$, *they uniquely determine a monic matrix polynomial* $P(X) = X^m + A_1X^{m-1} + \cdots + A_m$, *such that* $P(X_i) = Y_i$ *for* $i = 1,\cdots,m$, *if and only if* $V(X_1,\cdots,X_m)$ *is nonsingular.*

Proof: Let $\hat{Y}_i = Y_i - X_i^m$ and apply Theorem 3.1 to $(X_i, \hat{Y}_i)$. #

Let $M(X)$ have a complete set of solvents, $S_1,\cdots,S_m$, such that $V(S_1,\cdots,S_m)$ is nonsingular. According to Theorem 3.1, there exists a unique matrix polynomial

$$M_i(X) \equiv A_1^{(i)}X^{m-1} + \cdots + A_m^{(i)} \tag{3.1}$$

such that

$$M_i(S_j) = \delta_{ij}I. \tag{3.2}$$

Note that $M_i(X)$ has the same solvents as $M(X)$, except S_i has been deflated out. The $M_i(X)$'s are the fundamental matrix polynomials.

Denote by $V(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_m)$ the block Vandermonde at the m-1 solvents, $S_1,\cdots,S_m$, with S_i deleted.

Theorem 3.2 <u>If matrices</u> $S_1,\cdots,S_m$ <u>are such that</u> $V(S_1,\cdots,S_m)$ <u>is nonsingular, then there exist unique matrix polynomials</u> $M_i(X) \equiv A_1^{(i)}X^{m-1} + \cdots + A_m^{(i)}$, <u>for</u> $i = 1,\cdots,m$, <u>such that</u> $M_1(X),\cdots,M_m(X)$ <u>are fundamental matrix polynomials.</u> <u>If, furthermore,</u> $V(S_1,\cdots,S_{k-1},S_{k+1},\cdots,S_m)$ <u>is nonsingular, then</u> $A_1^{(k)}$ <u>is nonsingular.</u>

Proof: $V(S_1,\cdots,S_m)$ nonsingular implies that there exists a unique set of fundamental matrix polynomials, $M_1(X),\cdots,M_m(X)$. $V(S_1,\cdots,S_{k-1},S_{k+1},\cdots,S_m)$ nonsingular and Corollary 3.1 imply that there exists a unique monic matrix polynomial $N_k(X) \equiv X^{m-1} + N_1^{(k)}X^{m-2} + \cdots + N_m^{(k)}$, such that $N_k(S_j) = \underline{0}$ for $j \neq k$. Consider $Q_k(X) \equiv N_k(S_k)M_k(X)$. $Q_k(S_j) = N_k(S_j)$ for $j = 1,\cdots,m$. Since $V(S_1,\cdots,S_m)$ is nonsingular and both $Q_k(X)$ and $N_k(X)$ are of degree m-1, it follows that $Q_k(X) \equiv N_k(X)$. Thus, $N_k(X) \equiv N_k(S_k)M_k(X)$. Equating leading coefficients, we get $I = N_k(S_k)A_1^{(k)}$ and thus $A_1^{(k)}$ is nonsingular. #

3.2 <u>Representation Theorems.</u> The fundamental matrix polynomials, $M_1(X),\cdots,M_m(X)$, can be used in a generalized Lagrange interpolation formula. Paralleling the scalar case we get the following representation theorems.

Theorem 3.3 *If matrices* $S_1,\cdots,S_m$ *are such that* $V(S_1,\cdots,S_m)$ *is nonsingular, and* $M_1(S),\cdots,M_m(X)$ *are a set of fundamental matrix polynomials, then, for an arbitrary*

$$G(X) \equiv B_1X^{m-1} + \cdots + B_m, \tag{3.3}$$

it follows that

$$G(X) = \sum_{i=1}^{m} G(S_i)M_i(X). \tag{3.4}$$

Proof: Let $Q(X) = \sum_{i=1}^{m} G(S_i)M_i(X)$. Then $Q(S_i) = G(S_i)$ for $i = 1,\cdots,m$. Since the block Vandermonde is nonsingular, it follows that $Q(X)$ is unique and, hence, $G(X) \equiv Q(X)$. #

A lambda-matrix was defined as a matrix polynomial whose variable was restricted to the scalar matrix λI. Thus, the previous theorem holds for lambda-matrices as well.

Corollary 3.2 *Under the same assumptions as in Theorem* 3.3, *for an arbitrary lambda-matrix*

$$G(\lambda) \equiv B_1\lambda^{m-1} + \cdots + B_m, \tag{3.5}$$

it follows that

$$G(\lambda) = \sum_{i=1}^{m} G(S_i)M_i(\lambda). \tag{3.6}$$

A basis for lambda-vectors will be presented next.

Theorem 3.4 *If $M(\lambda)$ has distinct latent roots, $\rho_1,\cdots,\rho_{mn}$, with right latent vectors $b_1,\cdots,b_{mn}$, then for an arbitrary lambda-vector*

$$g(\lambda) \equiv v_1\lambda^{m-1} + \cdots + v_m \tag{3.7}$$

there exists a unique set of constants $\alpha_1,\ldots,\alpha_{mn}$, such that

$$g(\lambda) = \sum_{i=1}^{mn} \alpha_i \frac{M(\lambda)}{\lambda - \rho_i} b_i. \tag{3.8}$$

Proof: If the latent roots of $M(\lambda)$ are distinct, then the eigenvectors of the block companion matrix (Theorem 2.3 (iii)) form a basis for vectors of dimension mn. By equation (2.13) lambda-vectors $\frac{M(\lambda)}{\lambda - \rho_i} b_i$ are formed by partitioning the eigenvectors of the block companion matrix into the vector coefficients. The α_i's are those required to write $(v_1,\cdots,v_m)^T$ as a linear combination of the eigenvectors of the block companion matrix. #

3.3 Fundamental Matrix Polynomials. Fundamental matrix polynomials were defined such that $M_i(S_j) = \delta_{ij}I$. A result similar to equation (2.9) can be derived based on the fundamental matrix polynomials. It was previously (Section 2.1) developed using matrix polynomial division.

Theorem 3.5 *If $M(X)$ has a complete set of right solvents, $S_1,\cdots,S_m$, such that $V(S_1,\cdots,S_m)$ and $V(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_m)$ are nonsingular and $M_1(X),\cdots,M_m(X)$ are the set of fundamental matrix polynomials, then*

$$M_i(X)X - S_iM_i(X) \equiv A_1^{(i)}M(X), \qquad (3.9)$$

where $A_1^{(i)}$ is the leading matrix coefficient of $M_i(X)$.

Proof: Let $Q_i(X) \equiv M_i(X)X - S_iM_i(X)$. Note that $Q_i(S_j) = \underline{0}$ for all j. $M(X)$ is the unique monic matrix polynomial with right solvents $S_1,\cdots,S_m$ since $V(S_1,\cdots,S_m)$ is nonsingular. The leading matrix coefficient of $Q_i(X)$ is $A_1^{(i)}$ which is nonsingular, since $V(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_m)$ is nonsingular. Thus, $M(X) \equiv A_1^{(i)^{-1}} Q_i(X)$. #

A previous result (equation (2.5)) stated that if R_i was a left solvent of $M(X)$, then there exists a unique, monic polynomial $F_i(X)$ of degree m-1, such that

$$M(X) \equiv F_i(X)X - R_iF_i(X). \qquad (3.10)$$

Comparing equations (3.9) and (3.10), we obtain the following result.

Corollary 3.3 <u>Under the conditions of Theorem</u> 3.5 $F_i(X) \equiv \left[A_1^{(i)}\right]^{-1} N_i(X)$ <u>and</u>

$$R_i = \left[A_1^{(i)}\right]^{-1} S_i A_1^{(i)} \tag{3.11}$$

<u>is a left solvent of</u> $M(X)$.

If $M(X)$ has a complete set of right solvents, $S_1,\cdots,S_m$, such that $V(S_1,\cdots,S_m)$ and $V(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_m)$ for $i = 1,\cdots,m$ are all nonsingular, then, by equation (3.11), there exists a complete set of left solvents of $M(X)$, $R_1,\cdots,R_m$, such that R_i is similar to S_i for all i.

Corollary 3.4 <u>Under the conditions of Theorem</u> 3.5, <u>if</u> R_i <u>is defined as in equation</u> (3.11), <u>then</u>

$$\overline{M}_i(\lambda) \equiv \left[A_1^{(i)}\right]^{-1} N_i(\lambda) = (I\lambda - R_i)^{-1} M(\lambda). \tag{3.12}$$

Proof: The result follows from equation (2.11) and Corollary 3.3. #

CHAPTER 4

The Block Vandermonde

The block Vandermonde matrix is of fundamental importance to this work. This chapter considers the properties of the block Vandermonde.

It is well known that in the scalar case ($n = 1$),

$$\det V(s_1,\cdots,s_m) = \prod_{i>j} (s_i - s_j) \tag{4.1}$$

and, thus, the Vandermonde is nonsingular if the set of s_i's are distinct. One might expect that if the eigenvalues of X_1 and X_2 are disjoint and distinct, then $V(X_1,X_2)$ is nonsingular. That this is not the case is shown by the following example.

The determinant of the block Vandermonde at two points is

$$\det V(X_1,X_2) = \det \begin{pmatrix} I & I \\ X_1 & X_2 \end{pmatrix} = \det (X_2 - X_1). \tag{4.2}$$

Even if X_1 and X_2 have no eigenvalues in common, $X_2 - X_1$ may still be singular. The example $X_1 = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix}$ and $X_2 = \begin{pmatrix} 4 & 2 \\ 0 & 3 \end{pmatrix}$ yields $X_2 - X_1$ singular.

It will be shown that the X_1 and X_2 in this example cannot be the complete set of solvents of a monic matrix polynomial. First, however, the following is needed.

Lemma 4.1 Let matrix A have distinct eigenvalues, and N be a subspace of E^n of dimension d. Suppose further that if $v \in N$, then $Av \in N$. Under these conditions, d of the eigenvectors of A are in N.

Proof: Let $Av_i = \lambda_i v_i$ for $i = 1,\cdots,n$. The set of v_i's is a basis for E^n, since A has distinct eigenvalues. Let $v \in N \subset E^n$, and order the $\{v_i\}$ such that $v = \sum_{i=1}^{s} c_i v_i$ with $c_i \neq 0$ for $i = 1,\cdots,s$. Let $P(t) = \prod_{j=2}^{s} (t-\lambda_j)$, then $P(A)v_j = \underset{\sim}{0}$ for $j = 2,\cdots,s$. Hence,

$$P(A)v = \sum_{i=1}^{s} c_i P(A) v_i = c_1 P(A) v_1$$

$$= c_1 \left(\prod_{j=2}^{s} (A-\lambda_j I) \right) v_1$$

$$= c_1 \left(\prod_{j=2}^{s} (\lambda_1-\lambda_j) \right) v_1 .$$

Let $d_1 = c_1\left(\prod_{j=2}^{s} (\lambda_1-\lambda_j)\right) \neq 0$. Thus,

$v_1 = \frac{1}{d_1} P(A)v \in N$. Similarly, $v_i \in N$ for $i = 1,\cdots,s$. The lemma follows, since $v \in N$ was arbitrary. #

Theorem 4.1 <u>If $M(\lambda)$ has distinct latent roots, then there exists a complete set of right solvents of $M(X)$, $S_1,\cdots,S_m$, and for any such set of solvents, $V(S_1,\cdots,S_m)$ is nonsingular.</u>

Proof: The existence was proved in Theorem 2.7. $S_1,\cdots,S_m$, being right solvents of $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$, is equivalent to

$$(A_m,\cdots,A_1)\begin{pmatrix} I & \cdots & I \\ S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_1^{m-1} & \cdots & S_m^{m-1} \end{pmatrix} = \left(-S_1^m,\cdots,-S_m^m\right). \tag{4.3}$$

Assume $\det V(S_1,\cdots,S_m) = 0$, and let N be the null space of $V(S_1,\cdots,S_m)$. That is, $v \in N$ if and only if $V(S_1,\cdots,S_m)v = \underline{0}$. Since $A_1,\cdots,A_m$ in equation (4.3) exist, joining any row of $(-S_1,\cdots,-S_m)$ onto $V(S_1,\cdots,S_m)$ gives a larger matrix but with the same rank as $V(S_1,\cdots,S_m)$. Thus, for all $v \in N$, $\left(S_1^m,\cdots,S_m^m\right)v = \underline{0}$. Hence, for all $v \in N$

$$\underline{0} = \begin{pmatrix} S_1 & \cdots & S_m \\ S_1^2 & \cdots & S_m^2 \\ \vdots & & \vdots \\ S_1^m & \cdots & S_m^m \end{pmatrix} v = V(S_1,\cdots,S_m)\operatorname{diag}(S_1,\cdots,S_m)v. \tag{4.4}$$

Letting $A = \operatorname{diag}(S_1,\cdots,S_m)$, equation (4.4) shows that for all $v \in N$, $Av \in N$. Since A has distinct eigenvalues, Lemma 4.1 applies, and there are as many eigenvectors of A in N as the dimension of N. The eigenvalues of $\operatorname{diag}(S_1,\cdots,S_m)$ are the eigenvalues of the S_i's, and the eigenvectors are of the form $(0^T,v^T,0^T)$, where v is an eigenvector of one of the S_i's. This is because if

$$\begin{pmatrix} S_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & S_m \end{pmatrix} \begin{pmatrix} \underline{0} \\ v \\ \underline{0} \\ v \\ \underline{0} \end{pmatrix} = \lambda \begin{pmatrix} \underline{0} \\ v \\ \underline{0} \\ v \\ \underline{0} \end{pmatrix},$$

then $S_i v = \lambda v$ and $S_j w = \lambda w$. This cannot be since S_i and S_j do not have any common eigenvalues. Let an arbitrary eigenvector of $\operatorname{diag}(S_1,\cdots,S_m)$, $\left(0^T,v^T,0^T\right)^T$, be in N. Then

$$\begin{pmatrix} I & \cdots & I \\ S_1 & \cdots & S_m \\ \vdots & & \vdots \\ S_1^{m-1} & \cdots & S_m^{m-1} \end{pmatrix} \begin{pmatrix} \underline{0} \\ v \\ \underline{0} \end{pmatrix} = \underline{0}.$$

But then, $Iv = \underline{0}$ which is a contradiction. Thus $\det V(S_1,\cdots,S_m) \neq 0$. #

The example considered before this theorem was a case where matrices X_1 and X_2 had distinct and disjoint eigenvalues and $\det V(X_1,X_2) = 0$. Thus, by the theorem, they could not be a complete set of right solvents for a monic, quadratic matrix polynomial. In contrast with the theory of scalar polynomials, we have the following result.

Corollary 4.1 <u>There exist sets containing</u> m <u>matrices which are not a set of right solvents for any matrix polynomial of degree</u> m.

A generalization of equation (4.1), that the Vandermonde of scalars is the product of the differences of the scalars, will be given. Let $M^{(d)}_{S_1\cdots S_k}(X)$ be a monic matrix polynomial of degree $d \geq k$ with right solvents $S_1,\cdots,S_k$. The superscript d will be omitted if $d = k$. Note that this matrix polynomial need not necessarily exist, nor be unique.

Theorem 4.2 <u>If</u> $V(S_1,\cdots,S_k)$ <u>is nonsingular for</u> $k = 2,\cdots,r-1$, <u>then</u>

$$\det V(S_1,\cdots,S_r) = \det V(S_1,\cdots,S_{r-1}) \det M_{S_1\cdots S_{r-1}}(S_r). \tag{4.5}$$

Proof: The nonsingularity of $V(S_1,\cdots,S_{r-1})$ and Corollary 3.1 guarantee that $M_{S_1\cdots S_{r-1}}(X)$ exists uniquely. The determinant of $V(S_1,\cdots,S_r)$ will be evaluated by block Gaussian elimination using the fact that

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det\begin{pmatrix} A+EC & B+ED \\ C & D \end{pmatrix}. \tag{4.6}$$

$$\operatorname{Det} V(S_1,\cdots,S_r) = \det\begin{vmatrix} I & \cdots & I \\ S_1 & \cdots & S_r \\ \vdots & & \vdots \\ S_1^{r-1} & \cdots & S_r^{r-1} \end{vmatrix}$$

$$= \det\begin{pmatrix} I & I & \cdots & I \\ & S_2-S_1 & \cdots & S_r-S_1 \\ & \vdots & & \vdots \\ & S_2^{r-1}-S_1^{r-1} & \cdots & S_r^{r-1}-S_1^{r-1} \end{pmatrix}$$

$$= \det\begin{pmatrix} I & I & I & \cdots & I \\ & S_2-S_1 & S_3-S_1 & \cdots & S_r-S_1 \\ & & M^{(2)}_{S_1S_2}(S_3) & \cdots & M^{(2)}_{S_1S_2}(S_r) \\ & & \vdots & & \vdots \\ & & M^{(r-1)}_{S_1S_2}(S_3) & \cdots & M^{(r-1)}_{S_1S_2}(S_r) \end{pmatrix} \tag{4.7}$$

where $M^{(d)}_{S_1S_2}(X) = \left(X^d - S_1^d\right) - \left(S_2^d - S_1^d\right)(S_2-S_1)^{-1}(X-S_1)$. (S_2-S_1) is nonsingular, since $\det(S_2-S_1) = \det V(S_1,S_2) \neq 0$. It will be shown that after k steps of the block Gaussian elimination, the general term for the i,j block, $i,j > k$, is $M^{(i-1)}_{S_1\cdots S_k}(S_j)$. Assume it is true after k-1 steps. Then, after k steps, the i,j element is

$$M^{(i-1)}_{S_1\cdots S_{k-1}}(S_j) - M^{(i-1)}_{S_1\cdots S_{k-1}}(S_k)M^{(k-1)}_{S_1\cdots S_{k-1}}(S_k)^{-1}M^{(k-1)}_{S_1\cdots S_{k-1}}(S_j).$$

This is merely $M^{(i-1)}_{S_1\cdots S_k}(X)$ evaluated at $X = S_j$. Using the fact that the determinant of a block triangular matrix is the product of the determinants of the diagonal matrices, (see Householder [5]), the result follows. #

Corollary 4.2 <u>If</u> $V(S_1,\ldots,S_{k-1})$ <u>is nonsingular and</u> S_k <u>is not a weak solvent of</u> $M_{S_1\cdots S_{k-1}}(X)$, <u>then</u> $V(S_1,\cdots,S_k)$ <u>is nonsingular</u>.

It is useful to be able to construct matrix polynomials with a given set of right solvents.

Corollary 4.3 <u>Given matrices</u> $S_1,\cdots,S_m$ <u>such that</u> $V(S_1,\cdots,S_k)$ <u>is nonsingular for</u> $K = 2,\cdots,m$, <u>the iteration</u> $N_0(X) = I$

$$N_i(X) = N_{i-1}(X)X - N_{i-1}(S_i)S_iN^{-1}_{i-1}(S_i)N_{i-1}(X) \qquad (4.8)$$

<u>is defined and yields an</u> m <u>degree monic matrix polynomial</u> $N_m(X)$, <u>such that</u> $N_m(S_i) = \underline{0}$ <u>for</u> $i = 1,\cdots,m$.

Proof: $N_1(X) \equiv X - S_1 = M_{S_1}(X)$. Assume $N_k(X) \equiv M_{S_1\cdots S_k}(X)$. Then, from equation (4.8), $N_{k+1}(S_i) = \underline{0}$ for $i = 1,\cdots,k+1$ and, hence, $N_{k+1}(X) \equiv M_{S_1\cdots S_{k+1}}(X)$. The sequence of block Vandermonde being nonsingular guarantees the nonsingularity of $N_{i-1}(S_i)$. #

Corollary 4.4 <u>If</u> $V(S_1,\cdots,S_k)$ <u>is nonsingular for</u> $k = 2,\cdots,m$, <u>then</u> $S_1,\cdots,S_m$ <u>are a complete set of right solvents for</u> $M_{S_1\cdots S_m}(X)$.

Proof: The result follows directly from Theorem 3.5, where we obtained

$$(I\lambda - S_i)M_i(\lambda) = A_1^{(i)}M(\lambda). \qquad \# \quad (4.9)$$

CHAPTER 5

A Matrix Polynomial Algorithm

This chapter presents the paper's main algorithm. It computes solvents and is a generalization of one of Traub's methods. Section 5.1 gives the algorithm. A global convergence theorem is presented in Section 5.2. Section 5.3 considers computational aspects of the algorithm and has a detailed flow-chart of the method.

5.1 A Generalization of Traub's Algorithm. The following algorithm for matrix polynomials, in the scalar case, reduces to Traub's scalar polynomial algorithm.

Algorithm 1 (i) Let $G_o(X) = I$ and generate matrix polynomials $G_n(X)$ by

$$G_{n+1}(X) = G_n(X)X - \alpha_1^n M(X), \qquad (5.1)$$

for $n = 0,1,\cdots,L-1$, where

$$G_n(X) \equiv \alpha_1^n X^{m-1} + \cdots + \alpha_m^n. \qquad (5.2)$$

Then, (ii) let $X_o = \left(\alpha_1^L\right)\left(\alpha_1^{L-1}\right)^{-1}$ and generate

$$X_{i+1} = G_L(X_i)G_{L-1}^{-1}(X_i). \qquad (5.3)$$

The algorithm has two stages. The first, a generalization of Sebastião e Silva's algorithm (see Householder [4]), generates a sequence of matrix polynomials. Equation (5.1) ensures that each of these matrix polynomials is of degree less than m, the degree of $M(X)$. Under suitable conditions $\overline{G}_n(X) \equiv \left(u_1^n\right)^{-1} G_n(X)$ will be shown (in the next section) to converge to $\overline{M}_1(X)$, a monic fundamental matrix polynomial.

The second stage generates a sequence of matrix iterates which will be shown (in the next section) to converge to a solvent. The point at which one switches from stage one to stage two, the value of L, will be considered in Section 5.3.

5.2 <u>The Convergence Theorem</u>. In the proofs that Bernoulli's method and Traub's scalar polynomial algorithms converge, the main property needed is that if ρ_1 is a dominant root, then $(\rho_i/\rho_1)^n \to 0$ as $n \to \infty$, for ρ_i any other root. To generalize this property to solvents, the following result is needed, the proof of which was provided by P. A. Businger of Bell Telephone Laboratories.

Definition 5.1 *Matrix A dominates matrix B if all the eigenvalues of A strictly dominate, in modulus, those of B.*

Lemma 5.1 *If matrix A dominates matrix B, then* $A^{-n}CB^n \to \underset{=}{0}$ *as* $n \to \infty$, *for any constant matrix C.*

Proof: For any $\epsilon > 0$, let

$$B = P_B(\varepsilon)\Big(J_B(\varepsilon)\Big)P_B(\varepsilon)^{-1}, \tag{5.4}$$

where

$$J_B(\varepsilon) = \begin{pmatrix} \lambda_B & & & \\ \varepsilon & \lambda_B & & \\ & \ddots & \ddots & \\ & & \varepsilon & \lambda_B \end{pmatrix}. \tag{5.5}$$

See Ortega and Rheinboldt [16, p. 43] for a discussion on this modified Jordan form. Then,

$$\|B^n\| \leq \|P_B(\varepsilon)\| \; \|P_B(\varepsilon)^{-1}\| \; (\varepsilon + \max|\lambda_B|), \tag{5.6}$$

where the norm is the infinity norm. Noting that

$$\begin{pmatrix} \lambda_A & & & \\ \varepsilon & \lambda_A & & \\ & \ddots & \ddots & \\ & & \varepsilon & \lambda_A \end{pmatrix}^{-1} = \begin{pmatrix} \lambda_A^{-1} & & & \\ -\varepsilon\lambda_A^{-2} & \lambda_A^{-1} & & \\ & \ddots & \ddots & \\ & & -\varepsilon\lambda_A^{-2} & \lambda_A^{-1} \end{pmatrix}, \tag{5.7}$$

the result

$$\|A^{-n}\| \leq \|P_A(\varepsilon)\| \, \|P_A(\varepsilon)^{-1}\| \left(\frac{\varepsilon}{\min|\lambda_A^2|} + \frac{1}{\min|\lambda_A|} \right) \tag{5.8}$$

is similarly obtained, where $P_A(\varepsilon)$ is defined as in equation (5.4). Combining equations (5.6) and (5.8) we get

$$\|A^{-n}CB^{n}\| \le k\left[(\varepsilon+\max|\lambda_B|)\left(\frac{\varepsilon}{\min|\lambda_A^2|}+\frac{1}{\min|\lambda_A|}\right)\right]^{n}, \tag{5.9}$$

where k, a function of ε, is independent of n. When $\varepsilon = 0$, the constant to the n^{th} power is less than one, since $\max|\lambda_B|/\min|\lambda_A| < 1$. By continuity, there exists an $\varepsilon > 0$ so that the constant is still less than one, and, hence, $\|A^{-n}CB^{n}\| \to 0$ as $n \to \infty$. #

We now give the convergence theorem for Algorithm 1.

Theorem 5.1 *If*

(i) $M(X)$ *has a complete set of solvents*, $S_1,\cdots,S_m$,

(ii) S_1 *is a dominant solvent, and*,

(iii) $V(S_1,\cdots,S_m)$ *and* $V(S_2,\cdots,S_m)$ *are nonsingular*,

then (i) $\bar{G}_n(X) \equiv \left(\alpha_1^n\right)^{-1} G_n(X) \to \bar{M}_1(X)$, *where* $\bar{M}_1(X)$ *is the unique monic form of the fundamental matrix polynomial such that* $M_1(S_j) = \delta_{1j}I$, *and*

(ii) *for* L *sufficiently large*, X_L *of* (5.3) *converges to* S_1.

Proof of part (i): From equation (5.1), the result

$$G_n(S_1) = G_0(S_1)S_1^n = S_1^n \tag{5.10}$$

follows. By Theorem 3.3 and equation (5.10), we get

$$G_n(X) = \sum_{i=1}^{m} G_n(S_i)M_i(X) = \sum_{i=1}^{m} S_i^n M_i(X), \qquad (5.11)$$

and, thus,

$$\alpha_1^n = \sum_{i=1}^{m} S_i^n A_1^{(i)}. \qquad (5.12)$$

S_1 and $A_1^{(1)}$ are nonsingular and, thus, there is an N such that for $n \geq N$, α_1^n must be nonsingular, since using Lemma (5.1) and equation (5.12)

$$\alpha_1^n\left(S_1^n A_1^{(1)}\right)^{-1} \to I \qquad (5.13)$$

as $n \to \infty$. Using equations (5.11) and (5.12) and Lemma (5.1), we get, for $n \geq N$,

$$\begin{aligned}\overline{G}_n(X) &\equiv \left(\alpha_1^n\right)^{-1} G_n(X) \\ &= \left(\sum_{i=1}^{m} S_i^n A_1^{(i)}\right)^{-1} S_1^n S_1^{-n} \left(\sum_{i=1}^{m} S_i^n M_i(X)\right)\end{aligned}$$

(5.14)
(cont'd)

$$= \left(\sum_{i=1}^{m} S_1^{-n} S_i^n A_i^{(1)}\right)^{-1} \left(\sum_{i=1}^{m} S_1^{-n} S_i^n M_i(X)\right)$$

$$\to \left(A_1^{(1)}\right)^{-1} M_1(X) \equiv \overline{M}_1(X), \tag{5.14}$$

by Lemma 5.1. #

We defer the proof of part (ii) of the theorem to first obtain some results which will be needed in the proof.

Corollary 5.1 Under the hypotheses of Theorem 5.1,

$$\left(\alpha_1^n\right)^{-1} \alpha_1^{n+1} \to R_1 \tag{5.15}$$

as $n \to \infty$, where R_1 is the dominant left solvent.

Proof: Modification of equation (5.14) and Corollary 3.3 yields $\left(\alpha_1^n\right)^{-1} \alpha_1^{n+1} \to \left(A_1^{(1)}\right)^{-1} S_1 A_1^{(1)} = R_1$ as $n \to \infty$. #

The following lemmas all use the same hypotheses as in Theorem 5.1. Let

$$\phi_L(X) \equiv G_L(X) G_{L-1}^{-1}(X). \tag{5.16}$$

Thus, stage two of Algorithm 1, equation (5.3) is

$$X_{i+1} = \phi_L(X_i). \tag{5.17}$$

In Lemma 5.2 we show that every right solvent is a fixed point of $\phi_L(X)$ for each L. Lemma 5.4 shows that $\phi_L(X)$ is defined for all X in some neighborhood of the dominant solvent. Lemma 5.6 gives the local convergence of the second stage of Algorithm 1. Finally, Lemma 5.7 says that stage one will yield a point in the locally convergent region (Lemma 5.6) of the dominant solvent. Stage one supplies a sufficiently accurate starting value for the locally convergent stage two and, hence, the overall algorithm is globally convergent. The proof of part (ii) of Theorem 5.1 then immediately follows.

Lemma 5.2 $\phi_L(S) = S$ <u>for all</u> L <u>and any right solvent</u> S.

Proof: The result follows from equation (5.10) and the fact that $G_o(X) = I$. #

Lemma 5.3 <u>There exists a nontrivial ball</u> B, <u>centered at</u> S_1, <u>such that for all</u> $X \in B$

$$\text{(i)} \qquad \|I-M_1(X)\| \leq K < 1, \tag{5.18}$$

<u>and</u>

$$\text{(ii)} \qquad \|M_j(X)\| \leq D, \qquad j \neq 1, \tag{5.19}$$

<u>for some</u> D <u>independent of</u> j.

Proof: A matrix polynomial is a continuous function of its matrix variable. The results thus follow from continuity and the facts that $M_1(S_1) = I$ and $M_j(S_1) = \underline{0}$ for $j \neq 1$. #

It follows from Lemma 5.3 that for all $X \in B$, $M_1(X)$ is nonsingular and

$$\|M_1^{-1}(X)\| \leq \frac{1}{1 - \|I-M_1(X)\|} . \tag{5.20}$$

Lemma 5.4 <u>If</u> $X \in B$, <u>then there exists an</u> L' <u>such that</u> $\Phi_L(X)$ <u>is defined for every</u> $L \geq L'$.

Proof: For $X \in B$, let

$$V_j(X) = M_j(X)M_1^{-1}(X) \tag{5.21}$$

and

$$W_L(X) = \sum_{j=2}^{m} S_1^{-L}S_j^L V_j(X). \tag{5.22}$$

Then,

$$G_{L-1}(X) = \sum_{j=1}^{m} S_j^{L-1}M_j(X)$$

$$= S_1^{L-1}\left(I + \sum_{j=2}^{m} S_1^{-(L-1)}S_j^{L-1}V_j(X)\right) M_1(X)$$

$$= S_1^{L-1}\left(I+W_{L-1}(X)\right)M_1(X). \tag{5.23}$$

Note that $W_L(X) \to \underline{0}$ as $L \to \infty$ uniformly for $X \in B$. This follows since

$$\|V_j(X)\| = \|M_j(X)M_1^{-1}(X)\| \leq \frac{D}{1-K} < \infty \qquad (5.24)$$

by Lemma 5.3. Thus, $I + W_L(X) \to I$ as $L \to \infty$ and, hence, $I + W_L(X)$ is invertible for large L. By equation (5.23), $G_{L-1}(X)$ is invertible for large L and the result follows. #

Lemma 5.5 _If_ $X \in B$, _then_

$$\left\| S_j^L V_j(X) S_1^{-L} \right\| \leq \tau\sigma^L \|M_j(X)\| \|M_1^{-1}(X)\| \leq \frac{D\tau\sigma^L}{1-K}, \qquad (5.25)$$

where $0 \leq \sigma < 1$, _and_ τ _is a constant independent of_ L _and_ X.

Proof: The result follows from equation (5.9), where $\sigma = \max|\lambda_{S_j}|/\min|\lambda_{S_1}| < 1$ for $j \neq 1$. #

Lemma 5.6 _If_ $X_o \in B$ _and_ L _is sufficiently large_, _then_ $X_i = \phi_L(X_{i-1}) \to S_1$.

Proof: Let $X \in B$ and $L \geq L'$ of Lemma 5.4. Set

$$E_L(X) = \phi_L(X) - S_1. \qquad (5.26)$$

Then, since

$$\phi_L(X) = G_L(X)G_{L-1}^{-1}(X)$$

$$= \left(\sum_{j=1}^{m} S_j^L V_j(X)\right)\left(\sum_{j=1}^{m} S_j^{L-1} V_j(X)\right)^{-1},$$

it follows that

$$E_L(X) \sum_{j=1}^{m} S_j^{L-1} V_j(X) = \sum_{j=2}^{m} (S_j - S_1) S_j^{L-1} V_j(X).$$

Let

$$T_{j,L}(X) = S_j^{L-1} V_j(X) S_1^{-(L-1)}. \qquad (5.27)$$

Thus, by Lemma 5.5,

$$\| T_{j,L}(X) \| \le \frac{D\tau\sigma^L}{1-K} \to 0$$

as $L \to \infty$. Choose L large enough so that

$$\sum_{j=2}^{m} \| T_{j,L}(X) \| \le F \le 1 \qquad (5.28)$$

for all $X \in B$. Then,

$$E_L(X) \left[I + \sum_{j=2}^{m} T_{j,L}(X) \right] = \sum_{j=2}^{m} (S_j - S_1) T_{j,L}(X)$$

gives, by equation (5.25),

$$\| E_L(X) \| \le \sum_{j=2}^{m} \frac{\| S_j - S_1 \| \tau\sigma^{L-1} \| M_j(X) \| \| M_1^{-1}(X) \|}{1 - F} \qquad (5.29)$$

for all $X \in B$. A matrix polynomial is continuously differentiable. Since $M_j(S_1) = \underline{0}$ for $j \neq 1$, the result

$$\|M_j(X)\| \leq t\|X-S_1\|, \tag{5.30}$$

where $j \neq 1$, $t = \sup_{\substack{X \in B \\ j \neq 1}} \|M_j'(X)\|$, follows from the mean value theorem. Finally,

$$\|\phi_L(X)-S_1\| \leq c\sigma^{L-1}\|X-S_1\| \tag{5.31}$$

for all $X \in B$, where

$$c = \mathrm{abs}\left(\frac{\sum_{j=2}^{m} \|S_j-S_1\|\tau t}{(1-F)(1-K)}\right) \tag{5.32}$$

The result follows from equation (5.31), since $0 \leq \sigma < 1$ and L can be taken large enough so that $c\sigma^{L-1} < 1$.

#

The preceding lemma gave convergence for the second stage of Algorithm 1 if $X_0 \in B$. The next lemma shows that X_0 is in B if the first stage is continued long enough.

Lemma 5.7 <u>For</u> L <u>sufficiently</u> <u>large</u>, $\left(\alpha_1^L\right)\left(\alpha_1^{L-1}\right) \in B$.

Proof: Noting that $\alpha_1^L = \sum_{j=1}^{m} S_j^L A_1^{(j)}$, a proof similar to that in Lemma 5.6 will yield

$$\left(\alpha_1^L\right)\left(\alpha_1^{L-1}\right)^{-1} \to S_1 \qquad (5.33)$$

as $L \to \infty$. #

The second part of Theorem 5.1 can now be easily proved using Lemmas 5.2 through 5.7.

Proof of Part (ii) of Theorem 5.1: For L sufficiently large, $X_c \in B$ by Lemma 5.7. Lemma 5.6 then shows that $X_i \to S_1$. #

Equation (5.31) reveals the rate of convergence.

Corollary 5.2 $\|\phi_L(X)-S_1\| \leq c\sigma^{L-1}\|X-S_1\|$ <u>for all</u> $X \in B$, <u>where</u> $0 \leq c < 1$.

This corollary shows that even though the second stage is only linearly convergent, the asymptotic error constant can be made as small as desired by increasing the number of iterations of the first stage. The asymptotic error constant for stage one will depend on $\sigma = \max|\lambda_{S_j}|/\min|\lambda_{S_1}| < 1$, while that of stage two can be significantly faster than stage one. This is the purpose of the second stage, for equation (5.33) shows that stage one can also yield S_1.

5.3 Computational Considerations. A computational difficulty in generating the sequence $G_n(X)$ in stage one,

$$G_{n+1}(X) = G_n(X)X - \alpha_1^n M(X), \tag{5.34}$$

is that the matrix coefficients of $G_n(X)$ will grow exponentially. This may be avoided by generating $\tilde{G}_n(X)$ by

$$K_{n+1}(X) = \tilde{G}_n(X)X - \tilde{\alpha}_1^n M(X) \tag{5.35}$$

and

$$\tilde{G}_{n+1}(X) = \begin{cases} \dfrac{K_{n+1}(X)}{\| K_1^{n+1} \|} & \text{if} \quad K_1^{n+1} \neq \underline{0} \\ \\ K_{n+1}(X) & \text{otherwise,} \end{cases} \tag{5.36}$$

where $\tilde{\alpha}_1^n$ and K_1^n are the lead matrix coefficients of $\tilde{G}_n(X)$ and $K_n(X)$, respectively. Then let

$$G_{L-1}(X) \equiv \tilde{G}_{L-1}(X) \tag{5.37}$$

and

$$G_L(X) = G_{L-1}(X)X - \alpha_1^{L-1} M(X). \tag{5.38}$$

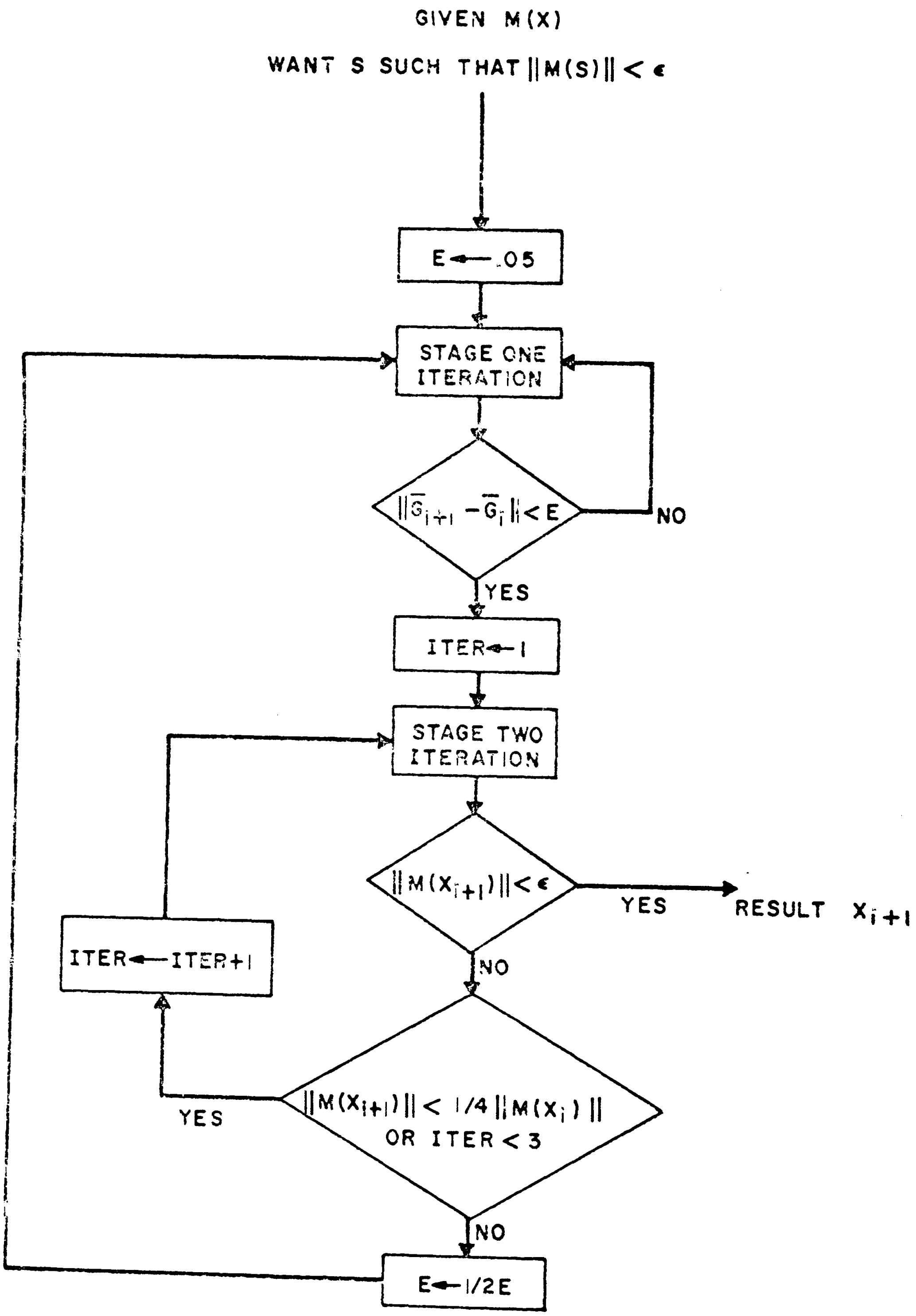
Now, $G_L(X)$ and $G_{L-1}(X)$ contain the same scalar constant that was built-up in normalizing $\tilde{G}_n(X)$ in equation (5.36). Thus, the constant vanishes in $\phi_L(X) = G_L(X)G_{L-1}^{-1}(X)$, and the growth of the coefficient has been stopped. Furthermore, $\bar{\bar{G}}_n(X) \equiv \bar{G}_n(X)$.

The following strategy is used to switch from stage one to stage two.

(i) Compute $\bar{G}_n(X)$ until the matrix polynomials tend to settle down.

(ii) Compute stage two, as long as rapid convergence appears to be occurring. If stage two is too slow or is diverging, resume stage one for several more steps.

A flow-chart of the algorithm that exhibits the strategy follows. It is guaranteed to work, using exact arithmetic, for any matrix polynomial satisfying the conditions of Theorem 5.1. The actual computer program that was used to test this algorithm appears in Appendix D.


GIVEN M(X)
WANT S SUCH THAT ||M(S)|| < ε
E ← .05
STAGE ONE ITERATION
$\|\bar{G}_{i+1} - \bar{G}_i\| < E$
NO
YES
ITER ← 1
STAGE TWO ITERATION
$\|M(X_{i+1})\| < \epsilon$
YES
RESULT X_{i+1}
NO
ITER ← ITER+1
$\|M(X_{i+1})\| < 1/4\,\|M(X_i)\|$ OR ITER < 3
YES
NO
E ← 1/2E

CHAPTER 6

The Block Bernoulli Method

This chapter covers a generalization of Bernoulli's scalar polynomial method to the matrix polynomial problem. A relationship is shown between it and Algorithm 1.

Definition 6.1 *For the matrix polynomial*

$$M(X) = X^m + A_1X^{m-1} + \cdots + A_m, \qquad (6.1)$$

the block Bernoulli iteration is

$$X_{i+1} + A_1X_i + \cdots + A_mX_{i-m+1} = \underline{0}, \qquad (6.2)$$

with $X_0, X_{-1}, \cdots, X_{-m+1}$ *given starting matrices.*

The general solution to the matrix difference equation (6.2) is obtained precisely as in the scalar case.

Theorem 6.1 *If* $S_1, \cdots, S_m$ *are right solvents of* $M(X)$, *such that* $V(S_1, \cdots, S_m)$ *is nonsingular, then*

$$X_i = S_1^i\alpha_1 + \cdots + S_m^i\alpha_m \qquad (6.3)$$

is the general solution to the matrix difference equation (6.2), *where* $\alpha_1, \cdots, \alpha_m$ *are matrices determined by the initial conditions.*

Proof: Substitution of equation (6.3) into equation (6.1) yields

$$\sum_{j=0}^{m} A_j X_{i+j+1} = \sum_{j=0}^{m} A_j \sum_{k=1}^{m} S_k^{i+j+1}\alpha_k$$

$$= \sum_{k=1}^{m} \left(\sum_{j=0}^{m} A_j S_k^{m-j} \right) S_k^{i+2j-m+1}\alpha_k = \underline{0}$$

where $A_0 = I$. The nonsingular block Vandermonde insures that $\alpha_1,\cdots,\alpha_m$ can be uniquely calculated in terms of $X_0, X_{-1},\cdots,X_{-m+1}$. If $\hat{X}_i$ is the general solution to equation (6.2) and $X_i = \hat{X}_i$ for the first m consecutive subscripts, then $X_i = \hat{X}_i$ for all i. #

In the scalar Bernoulli method, if there is a dominating root, then the ratio of the Bernoulli iterates converges to the root.

Theorem 6.2 *If $M(X)$ has solvents $S_1,\cdots,S_m$, such that S_1 is a dominant solvent, and $V(S_1,\cdots,S_m)$ is nonsingular, and if $X_0, X_{-1},\cdots,X_{-m+1}$ are chosen so that α_1 is nonsingular, then*

(i) $X_{n-1}^{-1}X_n \to \alpha_1^{-1}S_1\alpha_1$, *and*

(ii) $X_n X_{n-1}^{-1} \to S_1$ *as* $n \to \infty$.

Proof: Part (i) is obtained from

$$X_{n-1}^{-1}X_n = \left(\sum_{i=1}^{m} S_i^{n-1}\alpha_i\right)^{-1}\left(\sum_{i=1}^{m} S_i^{n}\alpha_i\right)$$

$$= \left(\alpha_1 + \sum_{i=2}^{m} S_1^{-(n-1)}S_i^{n-1}\alpha_i\right)^{1}\left(S_1\alpha_1 + \sum_{i=2}^{m} S_1^{-(n-1)}S_i^{n}\alpha_i\right)$$

$$\to \alpha_1^{-1}S_1\alpha_1.$$

For part (ii),

$$X_nX_{n-1}^{-1} = \left(\sum_{i=1}^{m} S_i^{n}\alpha_i\right)\left(\sum_{i=1}^{m} S_i^{n-1}\alpha_i\right)^{-1}$$

$$= \left(S_1 + W_nS_1^{n-1}\alpha_1^{-1}S_1^{-(n-1)}\right)\left(I + V_nS_1^{n-1}\alpha_1^{-1}S_1^{-(n-1)}\right)^{-1},$$

where

$$W_n = \sum_{j=2}^{m} S_j^{n}\alpha_j S_1^{-(n-1)} \tag{6.4}$$

and

$$V_n = \sum_{j=2}^{m} S_j^{n-1}\alpha_j S_1^{-(n-1)} . \tag{6.5}$$

Furthermore, $W_nS_1^{n-1}\alpha_1^{-1}S_1^{-(n-1)} \to \underline{0}$ and

$V_n S_1^{n-1} \alpha_1^{-1} S_1^{-(n-1)} \to \underline{0}$. Thus, $X_n X_{n-1}^{-1} \to S_1$. #

The block Bernoulli iteration (6.2) can also be written as

$$\begin{pmatrix} X_{i-m+2} \\ \vdots \\ X_i \\ X_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix} \begin{pmatrix} X_{i-m+1} \\ \vdots \\ X_{i-1} \\ X_i \end{pmatrix} \tag{6.6}$$

where X_i is a matrix of order n. Equation (6.6) looks like eigenvector powering except $\begin{pmatrix} X_{i-m+1} \\ \vdots \\ X_{i-1} \\ X_i \end{pmatrix}$ is not a vector in the usual sense. A theory of such power methods will be considered in Chapter 8.

Consider the same power-like method on the transpose of the matrix in equation (6.6). That is, consider

$$\begin{pmatrix} W_m^{i+1} \\ \vdots \\ W_2^{i+1} \\ W_1^{i+1} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & -A_m^T \\ I & & & -A_{m-1}^T \\ & \ddots & & \vdots \\ & & I & -A_1^T \end{pmatrix} \begin{pmatrix} W_m^i \\ \vdots \\ W_2^i \\ W_1^i \end{pmatrix} . \tag{6.7}$$

Multiplying out, the system

$$\begin{aligned}
W_m^{i+1} &= \qquad - A_m^T \; W_1^i \\
W_{m-1}^{i+1} &= W_m^i - A_{m-1}^T W_1^i \\
\vdots \quad & \quad \vdots \qquad \vdots \quad \vdots \\
W_1^{i+1} &= W_2^i - A_1^T \; W_1^i
\end{aligned} \tag{6.8}$$

results. Multiply the j^{th} equation on the left by $\left(X^T\right)^{j-1}$ and add. The result is

$$G_{i+1}(X) = G_i(X)X - \left(W_1^i\right)^T M(X), \tag{6.9}$$

where

$$G_i(X) \equiv \left(W_1^i\right)^T X^{m-1} + \cdots + \left(W_m^i\right)^T. \tag{6.10}$$

This is precisely stage one of Algorithm 1. These results are generalizations of what occurs in the scalar case. See Traub [21].

CHAPTER 7

A Lambda-Matrix Algorithm

In this chapter we present an algorithm, again based on Traub's scalar polynomial algorithm, to obtain a dominant latent root. Section 7.1 gives the algorithm and a convergence theorem. Section 7.2 considers another generalization of the Bernoulli method and its relationship to the algorithm of Section 7.1.

7.1 A Method Based on Lambda-Vectors. The basic approach to the lambda-matrix problem taken in this paper is to find a chain of solvents and, then, to find the eigenvalues of each matrix of the chain. For Algorithm 1 to yield a solvent, which is needed in this approach, a dominant solvent must exist. Since a dominant solvent need not exist, an alternative approach will be considered.

Algorithm 2 *Let* $g_o(\lambda)$ *be an arbitrary* m-1 *degree lambda-vector. Generate*

$$g_{k+1}(\lambda) = g_k(\lambda)\lambda - M(\lambda)v_1^{(k)}, \qquad (7.1)$$

where

$$g_k(\lambda) \equiv v_1^{(k)}\lambda^{m-1} + \cdots + v_m^{(k)}. \qquad (7.2)$$

Algorithm 2 is another generalization of Traub's scalar polynomial algorithm. It seeks a dominant latent root.

Theorem 7.1 <u>If</u>

(i) $M(\lambda)$ <u>has distinct latent roots</u>, $\rho_1, \cdots, \rho_{mn}$,

(ii) $|\rho_1| > |\rho_i|$ <u>for</u> $i \neq 1$, <u>and</u>

(iii) $r_1^T g_0(\rho_1) \neq 0$, <u>where</u> $r_1^T M(\rho_1) = \underline{0}^T$,

<u>then</u>

(i) $$\bar{g}_k(\lambda) \equiv \frac{g_k(\lambda)}{\max v_1^{(k)}} \to \frac{M(\lambda) b_1}{\lambda - \rho_1}, \quad \underline{\text{where}} \quad M(\rho_1) b_1 = \underline{0}$$

<u>and</u>

(ii) $$\frac{v_1^{(k+1)} - \rho_1 v_1^{(k)}}{\max v_1^{(k)}} \to \underline{0}.$$

Proof: By Theorem 3.4, the lambda-vector $g_n(\lambda)$ can be represented uniquely by

$$g_k(\lambda) = \sum_{i=1}^{mn} \beta_i^{(k)} \frac{M(\lambda)}{\lambda - \rho_i} b_i, \tag{7.3}$$

where $M(\rho_i) b_i = \underline{0}$. Thus,

$$v_1^{(k)} = \sum_{i=1}^{mn} \beta_i^{(k)} b_i. \tag{7.4}$$

Substituting equations (7.3) and (7.4) into equation (7.1), one gets $M(\lambda) \sum_{i=1}^{mn} \frac{\beta_i^{(k+1)} - \beta_i^{(k)} \rho_i}{\lambda - \rho_i} b_i = \underline{0}$

for all λ. Thus, $\beta_i^{(k)} = \beta_i \rho_i^k$, where $\beta_i = \beta_i^{(0)}$.
Using this,

$$\bar{g}_k(\lambda) = \frac{\sum_{i=1}^{mn} \beta_i \rho_i^k \frac{M(\lambda)}{\lambda - \rho_i} b_i}{\max \sum_{i=1}^{mn} \beta_i \rho_i^k b_i}$$

$$= \frac{M(\lambda) \sum_{i=1}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^k \frac{1}{\lambda - \rho_i} b_i}{\max \sum_{i=1}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^k b_i}$$

$$\rightarrow \frac{M(\lambda)}{\lambda - \rho_1} b_1$$

as $k \rightarrow \infty$, if $\beta_1 \neq 0$, since b_1 is unique to within a scalar multiple. Furthermore,
$g_o(\rho_1) = \beta_1 M'(\rho_1) b_1 + \sum_{i=2}^{mn} \frac{\beta_i M(\rho_1)}{\rho_1 - \rho_i} b_i$ and, thus,
since $r_1^T M(\rho_1) = \underline{0}^T$, we get

$$r_1^T g_o(\rho_1) = \beta_1 r_1^T M'(\rho_1) b_1. \qquad (7.5)$$

Finally, $r_1^T g_o(\rho_1) \neq 0$ implies $\beta_1 \neq 0$. For part (ii)

$$\frac{v_1^{(k+1)} - \rho_1 v_1^{(k)}}{\max v_1^{(k)}} = \frac{\sum_{i=2}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^k (\rho_i - \rho_1) b_i}{\max \sum_{i=1}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^k b_i} \to 0. \qquad \#$$

Let $(v)_r$ denote the r^{th} component of vector v.

Corollary 7.1 <u>Under the conditions of Theorem</u> 7.1, <u>if</u> $\left(b_1\right)_r \neq 0$, <u>then</u>, $\frac{\left(v_1^{(k+1)}\right)_r}{\left(v_1^{(k)}\right)_r} \to \rho_1$.

Proof:

$$\frac{\left(v_1^{(k+1)}\right)_r}{\left(v_1^{(k)}\right)_r} = \frac{\left(\sum_{i=1}^{mn} \beta_i \rho_i^{k+1} b_i\right)_r}{\left(\sum_{i=1}^{mn} \beta_i \rho_i^k b_i\right)_r}$$

$$= \rho_1 \frac{\left(\sum_{i=1}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^{k+1} b_i\right)_r}{\left(\sum_{i=1}^{mn} \beta_i \left(\frac{\rho_i}{\rho_1}\right)^k b_i\right)_r} \to \rho_1$$

as $k \to \infty$, as long as $(\beta_1 b_1)_r \neq 0$. #

If division of vectors is defined as componentwise division, then $v_1^{(k+1)} \div v_1^{(k)}$ is an n dimension vector, with each component an estimate of ρ_1. In a manner similar to the last two proofs, we get the following result.

Corollary 7.2 <u>Under the conditions of Theorem</u> 7.1,

$$\frac{v_1^{(k)}}{\max v_1^{(k)}} \to b_1.$$

Consider again, the first stage of Algorithm 1:

$$G_{k+1}(X) = G_k(X)X - \alpha_1^k M(X), \tag{7.6}$$

where

$$G_k(X) \equiv \alpha_1^k X^{m-1} + \cdots + \alpha_m^k. \tag{7.7}$$

Transpose both sides of equation (7.6) and substitute $X = \lambda I$ to get

$$G_{k+1}^T(\lambda) = G_k^T(\lambda)\lambda - M^T(\lambda)\left(\alpha_1^k\right)^T. \tag{7.8}$$

Let $g_k^i(\lambda)$ be the lambda-vector formed by taking the i^{th} column of the matrix coefficients of $G_k^T(\lambda)$. Then,

$$g_{k+1}^i(\lambda) = g_k^i(\lambda)\lambda - M^T(\lambda)v_{1,i}^k, \tag{7.9}$$

where $v_{1,i}^k$ is the leading vector coefficient of $g_k^i(\lambda)$.

Equation (7.9) is precisely Algorithm 2, operating on $M^T(\lambda)$. The latent roots of $M(\lambda)$ are the same as those of $M^T(\lambda)$. Thus, the computations of Algorithm 2 are done by Algorithm 1. Even if Algorithm 1 does not work, due to the lack of a dominant solvent, it is possible to obtain a dominant latent root by extracting the computations of Algorithm 2 from the computations (successful or not) of Algorithm 1.

The convergence theorem for Algorithm 2 has the requirement that $r_1^T g_o(\rho_1) \neq 0$. Since Algorithm 1 used $G_o(X) = I$, it follows that at least one column of equation (7.8) satisfies this requirement.

7.2 A Vector Bernoulli Method. A block (matrix) Bernoulli iteration was previously considered. Another generalization of Bernoulli's method is now presented. Similar ideas may be found in Guderley [3].

Definition 7.1 *For the lambda-matrix*

$$I\lambda^m + A_1\lambda^{m-1} + \cdots + A_m, \qquad (7.10)$$

the vector Bernoulli iteration is

$$v^{(k+1)} + A_1 v^{(k)} + \cdots + A_m v^{(k-m+1)} = \underline{0}, \qquad (7.11)$$

with $v^{(o)},\cdots,v^{(-m+1)}$ *given vectors.*

Equation (7.11) can be written as

$$\begin{pmatrix} v^{(k-m+2)} \\ \vdots \\ v^{(k)} \\ v^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix} \begin{pmatrix} v^{(k-m+1)} \\ \vdots \\ v^{(k-1)} \\ v^{(k)} \end{pmatrix} . \qquad (7.12)$$

This is just the eigenvector powering on the block transpose of the block companion matrix. Eigenvector powering on the block companion matrix is

$$\begin{pmatrix} v_m^{(k+1)} \\ \vdots \\ v_2^{(k+1)} \\ v_1^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & -A_m \\ I & & & -A_{m-1} \\ & \ddots & & \vdots \\ & & I & -A_1 \end{pmatrix} \begin{pmatrix} v_m^{(k)} \\ \vdots \\ v_2^{(k)} \\ v_1^{(k)} \end{pmatrix} . \qquad (7.13)$$

Multiplying out, we get

$$\begin{aligned} v_m^{(k+1)} &= \phantom{v_m^{(k)}} - A_m v_1^{(k)} \\ v_{m-1}^{(k+1)} &= v_m^{(k)} - A_{m-1} v_1^{(k)} \\ &\vdots \\ v_1^{(k+1)} &= v_2^{(k)} - A_1 v_1^{(k)} . \end{aligned} \qquad (7.14)$$

Then,

$$\varepsilon_{k+1}(\lambda) = \varepsilon_k(\lambda)\lambda - M(\lambda)v_1^{(k)}, \tag{7.15}$$

where the lambda-vector

$$\varepsilon_k(\lambda) \equiv v_1^{(k)}\lambda^{m-1} + \cdots + v_m^{(k)}, \tag{7.16}$$

is obtained by multiplying the i^{th} equation of (7.14) by λ^{i-1} and adding.

Equation (7.15) is precisely Algorithm 2. Consecutive substitutions of equations (7.14) yields

$$v_1^{(k+1)} + A_1v_1^{(k)} + \cdots + A_mv_1^{(k-m+1)} = \underline{0}. \tag{7.17}$$

Thus, the leading vector coefficient of Algorithm 2 is a vector Bernoulli iterate. This is a generalization of what occurs in Traub's [21] scalar polynomial algorithms.

CHAPTER 8

Block Eigenvalue Problem

A block eigenvalue problem is considered in this chapter. Let A be a given matrix of order mn. The matrix X of order n is desired such that there exists an mn by n matrix, V, of full rank, so that $AV = VX$. Power methods of the form $V_{i+1} = AV_i$ are considered, where V_i is an mn by n matrix. It was shown in Chapter 6 that the first stage of Algorithm 1 is of this form, where A is the block companion matrix. Sections 8.1 and 8.2 define the problem and consider complete sets of block eigenvalues. In Section 8.3 we present some generalizations of linear algebra with respect to this new formulation. The application of the new eigenvalue problem to the block companion matrix is given in Section 8.4. Also discussed is the relationship between block eigenvalues and right solvents. In Section 8.5 we present two algorithms based on eigenvector powering.

8.1 Block Eigenvectors. Let the term *block vector* denote an mn by n matrix that has been partitioned into a column of n by n blocks. It is equivalently an m-tuple, each of whose components is a square matrix.

Definition 8.1 *A matrix X of order n is a block eigenvalue of order n of matrix A of order mn, if there exists a block vector V of full rank, such that $AV = VX$. V is a block eigenvector of order n of A.*

Generally the order of a block eigenvalue or block eigenvector will be understood and will not be referred to explicitly.

A problem that has received a good deal of attention is that of finding a matrix X such that $AX = XB$, where matrices A and B, of orders m and n, respectively, are given. Jameson [6] and Gantmacher [2, p. 215] are amongst many authors who have considered this problem. The main result for this problem is that $AX = XB$ has only the trivial solution $X = \underline{0}$, if and only if A and B have no common eigenvalues. This result will be of use in this paper.

Returning to the block eigenvalue problem, we have the following.

Theorem 8.1 *If $AV = VX$ with V of full rank, then all the eigenvalues of X are eigenvalues of A.*

Proof: Let λ be an eigenvalue of X with eigenvector u. Thus, $AVu = VXu = \lambda Vu$. Therefore, either λ is an eigenvalue of A with eigenvector Vu or $Vu = \underline{0}$. Since V is an mn by n matrix and it is of full rank, there exists a left inverse to V. Thus, $Vu = \underline{0}$ can only occur if $u = \underline{0}$, which cannot happen since u is an eigenvector of X. #

Corollary 8.1 *If A is the block companion matrix, then all the eigenvalues of a block eigenvalue of A are latent roots of $M(\lambda)$.*

Proof: The result follows from Theorem 8.1 and the fact that the eigenvalues of the block companion matrix are latent roots of its associated lambda-matrix. #

8.2 Complete Sets of Block Eigenvalues. It will be shown that a solvent is a block eigenvalue of a block companion matrix. Furthermore, it will be proved that a matrix always has a block eigenvalue. Since a solvent does not always exist by Theorem 2.6, it follows that a block eigenvalue of a block companion matrix is not necessarily a solvent.

Definition 8.2 A set of block eigenvalues of a matrix is a complete set if the set of all the eigenvalues of these block eigenvalues is the set of eigenvalues of the matrix.

Theorem 8.2 Every matrix A, of order mn, has a complete set of block eigenvalues of order n.

Proof: Let $\rho_1,\cdots,\rho_n$ be any n eigenvalues of A and let $P_1,\cdots,P_n$ be their associated eigenvectors or principal vectors, where needed. Then, $V=(P_1,\cdots,P_n)$ is a block eigenvector with block eigenvalue in Jordan form. This process can be continued for each of the m sets of n eigenvalues of A. #

As an example of the construction in the above proof, let $\left(P_1P_2P_3P_4\right)^{-1} A(P_1P_2P_3P_4) = \begin{pmatrix} \rho & 1 & & \\ & \rho & 1 & \\ & & \rho & \\ & & & \mu \end{pmatrix}$. Then,

$A(P_1P_2) = (P_1P_2)\begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}$ and $A(P_1P_4) = (P_1P_4)\begin{pmatrix} \rho & 0 \\ 0 & \mu \end{pmatrix}$, and hence, $\begin{pmatrix} \rho & 0 \\ 0 & \mu \end{pmatrix}$ and $\begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}$ are a complete set of block eigenvalues of A.

Definition 8.3 In a complete set of block eigenvalues, one of them is weakly dominant, if all its eigenvalues are greater than or equal to the eigenvalues of any other block eigenvalue in the complete set.

The construction of Theorem 8.2 can be done such that the first block eigenvalue contains the n largest eigenvalues of the matrix. We thus get the following important result that was not true for solvents.

Corollary 8.2 Every block matrix has a complete set of block eigenvalues with one of them weakly dominant.

Block eigenvalues thus far considered have all been in Jordan form. However, unlike solvents, any matrix similar to a block eigenvalue is also a block eigenvalue. This follows, since, if $AV = VX$ and $Y = P^{-1}XP$, then $A(VP) = (VP)Y$, and VP is still of full rank.

8.3 Block Vector Algebra. We now consider some of the basic properties of block eigenvalues.

Definition 8.4 Block vectors, $V_1, \cdots, V_k$ of dimension mn by n, are block linearly independent, if $\sum_{i=1}^{k} V_iA_i = \underline{0}$ implies

<u>that</u> $A_i = \underline{0}$ <u>for all</u> i, <u>where</u> A_i <u>are matrices of order</u> n.

Note that a set of block vectors being block linearly dependent does not imply that one of them can be solved for as a combination of the others, since all the A_i's may be singular.

Lemma 8.1 <u>For</u> $i = 1,\cdots,m$, <u>let the block vector</u> $V_i = (v_{i1},\cdots,v_{in})$. <u>Then,</u> $V_1,\cdots,V_m$ <u>are block linearly independent if and only if</u> $\{v_{ij}\}$ <u>for</u> $i = 1,\cdots,m$, $j = 1,\cdots,n$, <u>are linearly independent in</u> E^{mn}.

Proof: (i) Assume $\{v_{ij}\}$ are linearly dependent. Thus, there exists $\{\alpha_{ij}\}$ not all zero, such that $\sum_{ij} \alpha_{ij} v_{ij} = \underline{0}$. Let A_i be a matrix whose first column is $(\alpha_{i1},\cdots,\alpha_{in})^T$, and the remainder of the matrix is zero. Then $\sum_{i=1}^{m} V_i A_i = \underline{0}$ and not all the $A_i = \underline{0}$.

(ii) Assume $\{V_i\}$ are block linearly dependent. Thus, there exists $\{A_i\}$ not all zero matrices, such that $\sum_{i=1}^{m} V_i A_i = \underline{0}$. Let k be such that there is an element in the k^{th} column of at least one A_i that is not zero. Then, $\sum_{ij} v_{ij}\left(A_i\right)_{jk} = \underline{0}$ since this is the k^{th} column of $\sum_{i=1}^{m} V_i A_i$ and, since, $\{v_{ij}\}$ are linearly dependent.

#

Definition 8.5 <u>Block vectors $V_1,\cdots,V_m$ of dimension mn by n form a block basis if for any V of the same dimension there exists a unique set of matrices $A_1,\cdots,A_m$ such that</u>
$$V = \sum_{i=1}^{m} V_i A_i .$$

Block vectors being block linearly independent and forming a block basis are related by the following.

Theorem 8.3 <u>Block vectors $V_1,\cdots,V_m$ of dimension mn by n form a block basis if and only if they are block linearly independent.</u>

Proof: Let V be a block vector of dimension mn by n.

$$V = \sum_{i=1}^{m} V_i A_i \quad \text{is equivalent to} \quad V = (V_1,\cdots,V_m)\begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}.$$

The matrix $(V_1,\cdots,V_m)$ is square and, by Lemma 8.1, nonsingular, if and only if $\{V_i\}$ are block linearly independent. #

A generalization of a matrix with distinct eigenvalues being similar to a diagonal matrix, is given by the next result.

Theorem 8.4 <u>If A has block eigenvalues $X_1,\cdots,X_m$ with block eigenvectors $V_1,\cdots,V_m$ that are block linearly independent, and if X is also a block eigenvalue of A, then X is a block eigenvalue of</u> $\operatorname{diag}(X_1,\cdots,X_m)$. <u>Furthermore,</u>

$$(V_1,\cdots,V_m)^{-1} A(V_1,\cdots,V_m) = \operatorname{diag}(X_1,\cdots,X_m). \qquad (8.1)$$

Proof: Equation (8.1) is easily verified. Let $AV = VX$. Then, by Theorem 8.3, there exists a unique set of n by n matrices, $\alpha_1,\cdots,\alpha_m$, such that
$V = \sum_{i=1}^{m} V_i\alpha_i$. Let $\Lambda = \left(\alpha_1^T,\cdots,\alpha_m^T\right)^T$. Thus,
$V = (V_1,\cdots,V_m)\Lambda$. Since $(V_1,\cdots,V_m)$ is nonsingular and V is of full rank, by definition, it follows that Λ is of full rank. Now, using equation (8.1), we get

$$
\begin{aligned}
(V_1,\cdots,V_m)\Lambda X = VX &= A(V_1,\cdots,V_m)\Lambda \\
&= (V_1,\cdots,V_m)\ \mathrm{diag}(X_1,\cdots,X_m)\Lambda .
\end{aligned}
$$

Finally, $\mathrm{diag}(X_1,\cdots,X_m)\Lambda = \Lambda X$ with Λ of full rank.

#

8.4 Block Companion Matrix. An application of the block eigenvalue problem is given below. We again consider the block companion matrix. Recall that

$$
C = \begin{pmatrix} 0 & \cdots & 0 & -A_m \\ I & & & -A_{m-1} \\ & \ddots & & \vdots \\ & & I & -A_1 \end{pmatrix} \qquad (8.2)
$$

and

$$C^B = \begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix} \qquad (8.3)$$

where

$$M(X) \equiv X^m + A_1 X^{m-1} + \cdots + A_m. \qquad (8.4)$$

It will be shown that a solvent is a block eigenvalue. The converse is not true, since a matrix similar to a block eigenvalue is also a block eigenvalue, but the same is not true of solvents.

The following is easily verified.

Theorem 8.5 _If_ S _is a right solvent of_ M(X), _then_ S _is a block eigenvalue of_ C^B _with block eigenvector_ $\begin{pmatrix} I \\ S \\ \vdots \\ S^{m-1} \end{pmatrix}$.

Unlike the scalar eigenvalue problem, the block eigenvalues, with respect to left and right block eigenvectors, are different.

Definition 8.6 _An_ n _by_ n _matrix_ Y _is a left block eigenvalue of dimension_ n _of_ A, _a matrix of order_ mn, _if there exists a block vector_ W _of dimension_ n _by_ mn _of full rank, such that_ WA = YW. W _is a left block eigenvector_.

A generalization of what occurs in the scalar case, (see Jenkins and Traub [8]), is given in the next theorem.

Theorem 8.6 <u>If</u> R <u>is a left solvent of</u> M(X), <u>then</u> R <u>is a left block eigenvalue of</u> C^B, <u>with left block eigenvector</u> $(D_{m-1},\cdots,D_1,I)$, <u>where</u>

$$D(\lambda) \equiv I\lambda^{m-1} + D_1\lambda^{m-2} + \cdots + D_{m-2}\lambda + D_{m-1} \equiv (I\lambda - R)^{-1}M(\lambda). \quad (8.5)$$

Proof: Let

$$(D_{m-1},\cdots,D_1,I)\begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix} = Y(D_{m-1},\cdots,D_1,I).$$

Multiplying out, we get

$$\begin{aligned} - A_m &= YD_{m-1} \\ D_{m-1} - A_{m-1} &= YD_{m-2}. \\ \vdots \qquad \vdots \quad &\quad \vdots \\ D_2 - A_2 &= YD_1 \\ D_1 - A_1 &= Y \end{aligned} \quad (8.6)$$

Consecutive substitutions yield
$Y^m + Y^{m-1}A_1 + \cdots + YA_{m-1} + A_m = \underline{0}$. Thus, $Y = R$, a left solvent of M(X). Now, multiply the i^{th}

equation of (8.6) by λ^{i-1}; add; let
$D(\lambda) \equiv I\lambda^{m-1} + D_1\lambda^{m-2} + \cdots + D_{m-1}$; and get
equation (8.5). #

In a similar manner, we find that if S is a right solvent of $M(X)$, then S is a block eigenvalue of C, with block eigenvector $\begin{pmatrix} V_{m-1} \\ \vdots \\ V_1 \\ I \end{pmatrix}$, where

$$M(\lambda)(I\lambda - S)^{-1} \equiv I\lambda^{m-1} + V_1\lambda^{m-2} + \cdots + V_{m-1}. \qquad (8.7)$$

Let R_i be a left solvent of $M(X)$. Then by equation (8.5) and Corollary 3.4, it follows that $\overline{M}_i(X) \equiv D_i(X)$, if the appropriate block Vandermondes are nonsingular. Also, by equation (3.12), $D_i(S_i) = \left(A_1^{(i)}\right)^{-1}$, which is the inverse of the leading matrix coefficient of the i^{th} fundamental matrix polynomial.

Let

$$V_i = \begin{pmatrix} I \\ S_i \\ \vdots \\ S_i^{m-1} \end{pmatrix} \qquad (8.8)$$

and

$$W_i = \left(D_{m-1}^{(i)}, \cdots, D_1^{(i)}, I\right), \tag{8.9}$$

where it is assumed that both $V(S_1, \cdots, S_m)$ and $V(S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_m)$ are nonsingular, and that $R_i = A_1^{(i)^{-1}} S_i A_1^{(i)}$ from equation (3.11).

The biorthogonality of right and left block eigenvectors is given by the following.

Theorem 8.7 <u>Under the above assumptions</u>

$$W_i V_j = \delta_{ij} A_1^{(i)^{-1}}. \tag{8.10}$$

Proof:

$$W_i V_j = \left(D_{m-1}^{(i)}, \cdots, D_1^{(i)}, I\right) \begin{pmatrix} I \\ S_j \\ \vdots \\ S_j^{m-1} \end{pmatrix}$$

$$= D_{m-1}^{(i)} + D_{m-2}^{(i)} S_j + \cdots + S_j^{m-1} = D_i(S_j)$$

$$= D_i(S_i) M_i(S_j) = \delta_{ij} D_i(S_i) = \delta_{ij} A_1^{(i)^{-1}}. \qquad \#$$

From Theorem 8.5 and Lemma 8.1 the result that $V(S_1, \cdots, S_m)$ is nonsingular, if and only if the block eigenvectors of C^B are block linearly independent, is easily obtained.

8.5 Algorithms for Block Eigenvectors. Consider now block powering methods, as in equations (6.6) and (6.7). Let

$(V)_k = V_k$, where $V = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{pmatrix}$ and V_i is an n by n matrix.

Algorithm 3 Let

$$U_{n+1} = AU_n\left(\left(AU_n\right)_k\right)^{-1}, \qquad (8.11)$$

where U_o is an arbitrary block vector of full rank and $1 \leq k \leq m$ is an arbitrary fixed integer.

The normalization in equation (8.11) depends upon the nonsingularity of $\left(AU_n\right)_k$.

Lemma 8.2 $U_n = A^n U_o\left(\left(A^n U_o\right)_k\right)^{-1}$.

Proof:

$$U_{n+1} = AU_n\left(\left(AU_n\right)_k\right)^{-1}$$

$$= A^2U_{n-1}\left(\left(AU_{n-1}\right)_k\right)^{-1}\left(\left(A^2U_{n-1}\left(\left(AU_{n-1}\right)_k\right)^{-1}\right)_k\right)^{-1}$$

$$= A^2U_{n-1}\left(\left(AU_{n-1}\right)_k\right)^{-1}\left(\left(A^2U_{n-1}\right)_k\left(\left(AU_{n-1}\right)_k\right)^{-1}\right)^{-1}$$

$$= A^2U_{n-1}\left(\left(A^2U_{n-1}\right)_k\right)^{-1} = \cdots = A^{n+1}U_o\left(\left(A^{n+1}U_o\right)_k\right)^{-1}. \#$$

With this identity, convergence can be proved.

Theorem 8.8 *Let $S_1,\cdots,S_m$ be a complete set of block eigenvalues of A with block eigenvectors $V_1,\cdots,V_m$. If S_1 dominates all the other block eigenvalues and U_o is in the span of $\{V_i\}$, that is $U_o = \sum_{i=1}^{m} V_i\alpha_i$, and α_1 is nonsingular, then $U_{n+1} = AU_n\left(\left(AU_n\right)_k\right)^{-1}$ converges to $V_1\left(\left(V_1\right)_k\right)^{-1}$, if $\left(V_1\right)_k$ is nonsingular.*

Proof:

$$U_n = \left(A^nU_o\right)\left(\left(A^nU_o\right)_k\right)^{-1}$$

$$= \left(\sum_{i=1}^{m} V_iS_i^n\alpha_i\right)\left(\left(\sum_{i=1}^{m} V_iS_i^n\alpha_i\right)_k\right)^{-1}$$

$$= \left(\sum_{i=1}^{m} V_iS_i^n\alpha_i\alpha_1^{-1}S_1^{-n}\right)\left(\sum_{i=1}^{m}\left(V_i\right)_k S_i^n\alpha_i\alpha_1^{-1}S_1^{-n}\right)^{-1}$$

$$\rightarrow V_1\left(\left(V_1\right)_k\right)^{-1},$$

as $n \rightarrow \infty$, by Lemma 5.1. Since, as shown above, $\left(AU_n\right)_k\alpha_1^{-1}S_1^{-n} \rightarrow \left(V_1\right)_k$, it follows that $\left(AU_n\right)_k$ is nonsingular for n sufficiently large since $\left(V_1\right)_k^{-1}$ exists by the hypothesis.

#

In the application to the block companion matrix, the existence of a k such that $\left(V_1\right)_k$ is nonsingular, is equivalent to the existence of a solvent. If a right solvent exists, k can be taken as 1 by Theorem 8.5. The converse is proved below.

Theorem 8.9 <u>If</u> $C^B V = VX$ <u>and</u> $(V)_1$ <u>is nonsingular</u>, <u>then</u> $S = (V)_1 X (V)_1^{-1}$ <u>is a right solvent</u>.

Proof:

Let $V(V)_1^{-1} = D = \begin{pmatrix} I \\ D_2 \\ \vdots \\ D_m \end{pmatrix}$. $V(V)_1^{-1}$ is a block eigenvector of C with block eigenvalue $S = (V)_1 X (V)_1^{-1}$. Thus,

$$\begin{pmatrix} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix} \begin{pmatrix} I \\ D_2 \\ \vdots \\ D_m \end{pmatrix} = \begin{pmatrix} I \\ D_2 \\ \vdots \\ D_m \end{pmatrix} S.$$

Multiplication yields $D_i = S^{i-1}$ and $D_m S + A_1 D_m + \cdots + A_m = \underline{0}$. Hence, S is a right solvent. #

Thus, Algorithm 3, applied to the block companion matrix, converges to a block eigenvector associated with a solvent. Since block eigenvalues always exist but solvents do not, it is necessary to consider a normalization which does not depend on the existence of solvents. A block eigenvalue yields, by Corollary 8.1, as much information to the latent root problem as a solvent does. The difficulty is that a deflation of the form $M(\lambda) = Q(\lambda)(I\lambda - S)$ is not available for block eigenvalues.

For a block vector V_j of full rank, let $\left(V_j\right)_{k_j}$ denote the n by n matrix formed by taking the first n rows of V_j that are linearly independent. Actually, the rule for choosing the n linearly independent rows is not important, as long as the rule yields a unique set of rows.

Algorithm 4 Let

$$U_{j+1} = AU_j\left(\left(AU_j\right)_{k_j}\right)^{-1}. \qquad (8.12)$$

If it is assumed that A is nonsingular and U_0 is of full rank, then AU_j will remain of full rank, and the iteration (8.12) will always be defined. It is the goal here to get U_j to converge to V_1, the block eigenvector corresponding to the dominant block eigenvalue of A. Since the dominant block eigenvalue cannot be singular, it follows that for U_j close to V_1, A is not required to be nonsingular to ensure that the normalization, (8.12), is defined.

Lemma 8.3 $U_j = A^j U_o\left(\left(A^j U_o\right)_{k_j}\right)^{-1}$.

Proof:

$$U_{j+1} = AU_j\left(\left(AU_j\right)_{k_j}\right)^{-1}$$

$$= A^2U_{j-1}\left(\left(AU_{j-1}\right)_{k_{j-1}}\right)^{-1}\left(\left(A^2U_{j-1}\left(\left(AU_{j-1}\right)_{k_{j-1}}\right)^{-1}\right)_{k_j}\right)^{-1}$$

$$= A^2U_{j-1}\left(\left(AU_{j-1}\right)_{k_{j-1}}\right)^{-1}\left(\left(A^2U_{j-1}\right)_{k_j}\left(\left(AU_{j-1}\right)_{k_{j-1}}\right)^{-1}\right)^{-1}$$

$$= A^2U_{j-1}\left(\left(A^2U_{j-1}\right)_{k_j}\right) = \cdots = A^{j+1}U_o\left(\left(A^{j+1}U_o\right)_{k_j}\right)^{-1}. \quad \#$$

Let $\left(V_1\right)_r$ denote the n by n matrix formed from the first n linearly independent rows of V_1. Convergence of Algorithm 4 can now be proved precisely, as in Theorem 8.8.

Theorem 8.10 Let $S_1,\cdots,S_m$ be a complete set of block eigenvalues of A with block eigenvectors $V_1,\cdots,V_M$. If S_1 dominates all the other block eigenvalues in the set and U_o is in the span of $\{V_i\}$, that is $U_o = \sum_{i=1}^{m} V_i\alpha_i$, and α_1 is nonsingular, then $U_{j+1} = AU_j\left(\left(AU_j\right)_{k_j}\right)^{-1}$ converges to $V_1\left(V_1\right)_r^{-1}$.

CHAPTER 9

Numerical Results

Eight numerical examples follow. All calculations were done on Cornell University's IBM 360/67 in APL. This is a time-sharing language that gives the numerical analyst flexibility in designing algorithms. It has complete matrix arithmetic and does all calculations in double precision.

9.1 Consider the monic cubic matrix polynomial

$$M(X) = X^3 + \begin{pmatrix} -6 & 6 \\ -3 & -15 \end{pmatrix} X^2 + \begin{pmatrix} 2 & -42 \\ 21 & 65 \end{pmatrix} X + \begin{pmatrix} 18 & 66 \\ -33 & -81 \end{pmatrix}.$$

Algorithm 1 yields for stage one

$$\bar{G}_0(X) = X^2,$$

$$\bar{G}_1(X) = X^2 + \begin{pmatrix} -1.444 & 2.222 \\ -1.111 & -4.778 \end{pmatrix} X + \begin{pmatrix} -0.667 & -4.667 \\ 2.333 & 6.333 \end{pmatrix},$$

$$\bar{G}_2(X) = X^2 + \begin{pmatrix} -1.821 & 2.979 \\ -1.490 & -6.290 \end{pmatrix} X + \begin{pmatrix} -1.105 & -6.865 \\ 3.432 & 9.192 \end{pmatrix},$$

$$\bar{G}_3(X) = X^2 + \begin{pmatrix} -1.956 & 3.356 \\ -1.678 & -6.989 \end{pmatrix} X + \begin{pmatrix} -1.394 & -8.061 \\ 4.030 & 10.697 \end{pmatrix},$$

$$\bar{G}_4(X) = X^2 + \begin{pmatrix} -2.008 & 3.574 \\ -1.787 & -7.368 \end{pmatrix} X + \begin{pmatrix} -1.586 & -8.762 \\ 4.381 & 11.557 \end{pmatrix},$$

and

$$\overline{G}_5(X) = X^2 + \begin{pmatrix} -2.026 & 3.711 \\ -1.856 & -7.593 \end{pmatrix} X + \begin{pmatrix} -1.715 & -9.193 \\ 4.597 & 12.075 \end{pmatrix},$$

and for stage two

$$X_o = \begin{pmatrix} 3.9925 & -2.4261 \\ 1.2131 & 7.6317 \end{pmatrix},$$

$$X_1 = \begin{pmatrix} 3.9729 & -2.0892 \\ 1.0446 & 7.1067 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} 3.9927 & -2.0179 \\ 1.0089 & 7.0195 \end{pmatrix},$$

$$X_3 = \begin{pmatrix} 3.9985 & -2.0034 \\ 1.0017 & 7.0035 \end{pmatrix},$$

$$X_4 = \begin{pmatrix} 3.9997 & -2.0006 \\ 1.0003 & 7.0006 \end{pmatrix},$$

and

$$X_5 = \begin{pmatrix} 3.9999 & -2.0001 \\ 1.0001 & 7.0001 \end{pmatrix}.$$

$S_1 = \begin{pmatrix} 4 & -2 \\ 1 & 7 \end{pmatrix}$ is a dominant right solvent of the matrix polynomial.

9.2 Consider the monic, cubic matrix polynomial having right solvents $S_1 = \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix}$, $S_2 = \begin{pmatrix} 5 & 1 \\ -2 & 2 \end{pmatrix}$ and $S_3 = \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}$, which have eigenvalues 5,6; 3,4 and 1,2, respectively. Thus, S_1 is a dominant solvent. Furthermore, $V(S_1,S_2,S_3)$ and $V(S_2,S_3)$ are nonsingular. The unique monic matrix polynomial having these solvents, which was obtained using Corollary 4.3, is

$$M(X) = X^3 + \begin{pmatrix} -11.79104478 & 0.82089552 \\ 1.91044776 & -9.20895522 \end{pmatrix} X^2$$

$$+ \begin{pmatrix} 42.34328358 & -10.16417910 \\ -13.43283582 & 25.64179104 \end{pmatrix} X$$

$$+ \begin{pmatrix} -50.35820896 & 21.88059701 \\ 19.58208955 & -22.80597015 \end{pmatrix} .$$

The corresponding lambda-matrix has latent roots and latent vectors

Root	Vector
1	$(1,1.5)^T$
2	$(1,1)^T$
3	$(1,-2)^T$
4	$(1,-1)^T$
5	$(1,-1)^T$
6	$(1,-.5)^T$

From these results, we find that $S_4 = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}$ is also a solvent. Its eigenvalues are 3 and 5 and, hence, it yields only redundant information for the lambda-matrix problem. Note that the only combination of latent roots that cannot be eigenvalues of a solvent are 4 and 5.

For this problem

$$\overline{M}_1(X) = X^2 + \begin{pmatrix} -5\ 7/9 & 1\ 4/9 \\ 1\ 8/9 & -4\ 2/9 \end{pmatrix} X + \begin{pmatrix} 8\ 7/9 & -4\ 1/9 \\ -3\ 8/9 & 4\ 5/9 \end{pmatrix},$$

to which $\overline{G}_n(X)$ is to converge. Letting $G_c(X) = X^2$, we get

$$\overline{G}_1(X) = X^2 + \begin{pmatrix} -3.541 & .678 \\ .724 & -2.644 \end{pmatrix} X + \begin{pmatrix} 4.183 & -1.708 \\ -1.259 & 2.122 \end{pmatrix},$$

$$\overline{G}_2(X) = X^2 + \begin{pmatrix} -5.696 & 1.407 \\ 1.759 & -4.161 \end{pmatrix} X + \begin{pmatrix} 8.566 & -3.986 \\ -3.553 & 4.357 \end{pmatrix},$$

and

$$\overline{G}_3(X) = X^2 + \begin{pmatrix} -5.770 & 1.441 \\ 1.876 & -4.216 \end{pmatrix} X + \begin{pmatrix} 8.756 & -4.099 \\ -3.854 & 4.535 \end{pmatrix}.$$

The ratio of the leading matrix coefficients, which is to converge to $S_1 = \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix}$, results in

$$\left(\alpha_1^1\right)\left(\alpha_1^0\right)^{-1} = \begin{pmatrix} 11.791 & -.821 \\ -1.910 & 9.209 \end{pmatrix},$$

$$\left(\alpha_1^{10}\right)\left(\alpha_1^9\right)^{-1} = \begin{pmatrix} 6.874 & 1.682 \\ -.877 & 4.308 \end{pmatrix},$$

and

$$\left(\alpha_1^{20}\right)\left(\alpha_1^{19}\right)^{-1} = \begin{pmatrix} 6.983 & 1.966 \\ -.983 & 4.034 \end{pmatrix}.$$

Algorithm 2 which yields a dominant latent root was shown to be obtainable from the first stage of Algorithm 1. The iteration for this problem is

	Latent Root Estimate
1	11.791044
2	8.332911
3	7.247455
4	6.743387
5	6.467439
6	6.302969
7	6.200093
8	6.133848
9	6.090399
10	6.061549
11	6.042225
12	6.029191
13	6.020346
14	6.014309
15	6.010162
16	6.007294
17	6.005296
18	6.003892
19	6.002895
20	6.002181
21	6.001663
22	6.001283
23	6.001000
24	6.000787
25	6.000626
26	6.000501
27	6.000404
28	6.000327
29	6.000267
30	6.000218

All of the iterations thus far described have been linearly convergent. The ratio of the errors has been .8, which is the ratio of the smallest eigenvalue of the dominant solvent and the largest of the next dominant solvent. The second stage should also be linear, but with a ratio of errors $C(.8)^{L-1}$. The results are

$$
\begin{array}{lccc}
 & \underline{L = 10} & \underline{L = 5} & \underline{L = 2} \\[6pt]
X_o = & \begin{pmatrix} 6.8738 & 1.6815 \\ -.8769 & 4.3084 \end{pmatrix} & \begin{pmatrix} 6.8632 & 1.0039 \\ -.8216 & 4.9284 \end{pmatrix} & \begin{pmatrix} 8.3287 & -.0258 \\ -1.2123 & 6.4868 \end{pmatrix} \\[10pt]
X_1 = & \begin{pmatrix} 6.9766 & 1.9515 \\ -.9770 & 4.0475 \end{pmatrix} & \begin{pmatrix} 6.8378 & 1.5718 \\ -.8495 & 4.3913 \end{pmatrix} & \begin{pmatrix} 7.0843 & .5261 \\ -.8895 & 5.3558 \end{pmatrix} \\[10pt]
X_2 = & \begin{pmatrix} 6.9963 & 1.9927 \\ -.9964 & 4.0072 \end{pmatrix} & \begin{pmatrix} 6.9153 & 1.8115 \\ -.9215 & 4.1741 \end{pmatrix} & \begin{pmatrix} 6.8201 & .8844 \\ -.8009 & 4.9547 \end{pmatrix} \\[10pt]
X_3 = & \begin{pmatrix} 6.9994 & 1.9989 \\ -.9995 & 4.0011 \end{pmatrix} & \begin{pmatrix} 6.9602 & 1.9165 \\ -.9530 & 4.0774 \end{pmatrix} & \begin{pmatrix} 6.7670 & 1.1458 \\ -.7878 & 4.7206 \end{pmatrix} \\[10pt]
X_4 = & \begin{pmatrix} 6.9999 & 1.9998 \\ -.9999 & 4.0002 \end{pmatrix} & \begin{pmatrix} 6.9819 & 1.9630 \\ -.9832 & 4.0343 \end{pmatrix} & \begin{pmatrix} 6.7763 & 1.3414 \\ -.8043 & 4.5559 \end{pmatrix} \\[10pt]
X_{15} = & \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix} & \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix} & \begin{pmatrix} 6.9790 & 1.9573 \\ -.9819 & 4.0367 \end{pmatrix}.
\end{array}
$$

The ratio of the errors, which by Corollary 5.2 should be $C(.8)^{L-1}$, was found for large values of i to be

L	$\lVert X_{i+1}-S_1 \rVert_\infty / \lVert X_i - S_1 \rVert_\infty$
10	.15
8	.23
6	.36
4	.54
2	.78
1	.91

This shows that by increasing the number of iterations of stage one, stage two can be made to converge more rapidly.

9.3 Consider the matrix polynomial

$$M(X) = X^2 + \begin{pmatrix} -11.44382802 & 3.420249653 \\ 0.8613037448 & -5.556171983 \end{pmatrix} X^2$$

$$+ \begin{pmatrix} 41.02912621 & -20.93481276 \\ 0.5533980583 & 7.332871012 \end{pmatrix} X$$

$$+ \begin{pmatrix} -39.65603329 & 23.56171983 \\ 0.6074895978 & -3.386962552 \end{pmatrix}.$$

It has a complete set of solvents, $S_1 = \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix}$ and $S_2 = S_3 = \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}$. The eigenvalues of S_1 are 5 and 6, while while the eigenvalues of S_2 are 1 and 2. Clearly, $V(S_1, S_2, S_3)$ and $V(S_2, S_3)$ are singular. Algorithm 1 converged for all values of L. With $L = 6$, we get

$$X_0 = \begin{pmatrix} 6.7783 & 1.2464 \\ -1.0231 & 3.9215 \end{pmatrix},$$

$$X_1 = \begin{pmatrix} 6.9896 & 1.9764 \\ -1.0011 & 3.9975 \end{pmatrix},$$

and

$$X_2 = \begin{pmatrix} 6.9997 & 1.9995 \\ -1.0000 & 3.9999 \end{pmatrix}.$$

The convergence is fast, though linear, since the asymptotic error constant is $(.4)^6$.

9.4 Consider the problem

$$M(X) = X^2 + \begin{pmatrix} -12.4 & 4.4 \\ 1.6 & -8.6 \end{pmatrix} X^2 + \begin{pmatrix} 52.6 & -29.2 \\ -10.4 & 22.8 \end{pmatrix} X + \begin{pmatrix} -73.2 & 40.8 \\ 16.8 & -19.2 \end{pmatrix}.$$

This problem has a complete set of solvents, $S_1 = \begin{pmatrix} 7 & 2 \\ -1 & 4 \end{pmatrix}$, $S_2 = \begin{pmatrix} 1 & 0 \\ -2 & 2 \end{pmatrix}$, and $S_3 = \begin{pmatrix} 3 & 2 \\ 0 & 4 \end{pmatrix}$. S_1 dominates, $V(S_1,S_2,S_3)$ is nonsingular, and $V(S_2,S_3)$ is singular. $M_1(X)$ exists uniquely, but its leading matrix coefficient is singular. Hence $\lim_{n\to\infty} \overline{G}_n(X)$ does not exist. However, Algorithm 1 converged. This is because the second stage needs the ratio of $G_L(X)$ and $G_{L-1}(X)$, not $\overline{G}_L(X)$. For this type of problem, the equation $X_o = \alpha_1^L\left(\alpha_1^{L-1}\right)^{-1}$ can cause difficulties because α_1^{L-1} can become singular. For this problem, however, the ratio did exist since α_1^{L-1} did not quite become singular. If it had, a random X_o would have been used. After twenty iterations of the first stage,

$$X_o = \begin{pmatrix} 5.0260 & -2.0376 \\ -.5065 & 5.0094 \end{pmatrix}.$$

Then,

$$X_1 = \begin{pmatrix} 5.1741 & -1.6544 \\ -.5435 & 4.9136 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} 6.6745 & 1.3489 \\ -.9186 & 4.1628 \end{pmatrix},$$

and

$$X_3 = \begin{pmatrix} 6.9929 & 1.9857 \\ -.9982 & 4.0036 \end{pmatrix}.$$

9.5 Consider the quadratic

$$M(X) = X^2 + \begin{pmatrix} 7 & 8 \\ 8 & 10 \end{pmatrix} X + \begin{pmatrix} 9 & 3 \\ 4 & 4 \end{pmatrix}.$$

The corresponding lambda-matrix has latent roots -16,05113, -.4215 and $-.2637 \pm 1.8649i$. There exist two solvents having these as their eigenvalues, but neither can dominate, since there is a complex pair of latent roots whose absolute value is between the two other latent roots. Algorithm 1 did not converge, but Algorithm 2, whose computations are done by Algorithm 1, did converge to yield the dominant latent root, -16.05113. The order of the matrix coefficients was then reversed and the minimum latent root was found. Using these results, a solvent was formed, deflated, and the new problem yielded a solvent with eigenvalues which were the remaining complete pair of latent roots. This problem suggests the use of a random complex shift of the variable in the lambda-matrix. This will break up troublesome complex pairs of latent roots. With a shift of i, Algorithm 1 converged with no difficulties. All computations were done in the complex domain.

9.6 Consider the quadratic

$$M(X) = X^2 + \begin{pmatrix} -1 & -6 \\ 2 & -9 \end{pmatrix} X + \begin{pmatrix} 0 & 12 \\ -2 & 14 \end{pmatrix}.$$

The corresponding lambda-matrix has latent roots 1,2,3,4 with corresponding latent vectors $(1,0)^T$, $(0,1)^T$, $(1,1)^T$, $(1,1)^T$. The problem has a complete set of solvents $S_1 = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ and $S_2 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$. Other solvents have eigenvalues 1,2; 1,4 and 2,3. The only pair which cannot be the eigenvalues of a solvent is 3,4. Thus, no dominant solvent exists and Algorithm 1 did not converge. However, Algorithm 2, as computed by Algorithm 1, yielded the dominant latent root, 4.

Reversing the order of matrix coefficients has the effect of making the latent roots the reciprocals of the original latent roots. The right solvents are the inverse of the original ones. Thus, 1 and ½ are the new dominant latent roots. Algorithm 1 converged to $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$, and, hence, the solvent $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ was found for the original problem. Note that for the problem for which Algorithm 1 did converge, there was no complete set of solvents which included the dominant solvent $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.

9.7 Lancaster considers a test problem which "depends on a parameter whose value determines the proximity of clustered roots" [13, p. 90]. Consider $M(X) = X^2 + A_1X + A_2$, where

$$A_1 = \begin{pmatrix} 3\alpha & -(1+\alpha^2+2\beta^2) & \alpha(1+2\beta^2) & -\beta^2(\alpha^2+\beta^2) \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} -1+2\alpha^2 & \alpha-\alpha(\alpha^2+2\beta^2) & 2\alpha^2\beta^2 & -\alpha\beta^2(\alpha^2+\beta^2) \\ 2\alpha & -(\alpha^2+2\beta^2) & 2\alpha\beta^2 & -\beta^2(\alpha^2+\beta^2) \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

where $\beta = \alpha + 1$. The eight latent roots of $M(\lambda)$ are

$$0\ ,$$

$$\pm i$$

$$-\alpha$$

$$\pm(1+\alpha)i\ ,$$

and $$-\alpha\pm(1+\alpha)i\ .$$

Algorithm 1 was tested, and worked for $\alpha = 2, 1, .5, .1$ and $.001$. When α is made small, the smallest eigenvalue of the dominant solvent approaches the largest eigenvalue of the next solvent. Thus, convergence is considerably slower for smaller α. Using the code in Appendix D, the results were

α	L	Iterations	$\|M(X_i)\|$
2	10	3	7×10^{-6}
1	10	2	9×10^{-9}
.5	10	2	8×10^{-6}
.1	28	7	2×10^{-6}
.001	30	6	.004

9.8 Finally, consider the intriguing problem

$$M(X) = X^2 + \begin{pmatrix} -1 & -1 \\ -2 & -2 \end{pmatrix} X + \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}.$$

Note that

$$M(\lambda) = \left(I\lambda - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)\left(I\lambda - \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}\right).$$

$S_1 = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}$ is a dominant solvent, but it can be shown that there is no corresponding S_2 that would form a complete set of solvents. Letting $G_0(X) = X$, we get

$$G_1(X) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} X - \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$$

and

$$G_2(X) = \begin{pmatrix} 1 & 1 \\ 6 & 6 \end{pmatrix} X - \begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix}.$$

With

$$\phi_2(X) = G_2(X) G_1^{-1}(X)$$

$$= \left[\begin{pmatrix} 1 & 1 \\ 6 & 6 \end{pmatrix} X - \begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix}\right]\left[\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} X - \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}\right]^{-1},$$

it is easily seen that $\phi_2(X) \equiv S_1$ for all X such that $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} X - \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$ is nonsingular. Thus, the exact solution is obtained in one iteration of stage two for any X satisfying this one easy condition.

APPENDIX A

Nonmonic Lambda-Matrices

This paper has considered only matrix polynomials (and lambda-matrices), where the identity matrix was the leading matrix coefficient. Consider now, the matrix polynomial

$$M(X) \equiv A_0 X^m + A_1 X^{m-1} + \cdots + A_m. \qquad (A.1)$$

If A_0 is nonsingular, then $\overline{M}(X) \equiv A_0^{-1} M(X)$ is the problem that is dealt with in the body of this paper. If R is a left solvent of M(X), the $\overline{R} = A_0^{-1} R A_0$ is a left solvent of $\overline{M}(X)$.

The case where A_0 is singular presents some difficulty in the matrix polynomial problem. Franklin [1] considers the problem $M(X) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^2 + \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix} X + \begin{pmatrix} 0 & 0 \\ 0 & 6 \end{pmatrix} = \underline{0}$, which has a solvent $\begin{pmatrix} 0 & -2 \\ a & b \end{pmatrix}$ for all values of $\underline{a}$ and $\underline{b}$. Thus, a matrix polynomial with both A_0 and A_m singular can have solutions with variable eigenvalues.

If A_m is nonsingular then

$$M^R(X) \equiv A_m X^m + A_{m-1} X^{m-1} + \cdots + A_0 \qquad (A.2)$$

can be used. The solvents of $M^R(X)$ are the inverses of the solvents of M(X). M(X) does not have any singular solvents since A_m is nonsingular. However, if $M^R(X)$ has a complete set of solvents, then one of them must be singular, since zero is a latent root of $M^R(X)$. This follows since $\det M^R(0) = \det A_0 = 0$.

In contrast to the matrix polynomial problem, the latent roots of the lambda-matrix problem

$$M(\lambda) \equiv A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_m \qquad \text{(A.3)}$$

can be calculated, even if A_0 is singular. If A_m is singular, then $\lambda = 0$ is a latent root of $M(\lambda)$. If c is not a latent root of $M(\lambda)$, then $A_m(c)$ is nonsingular, where

$$M_c(\lambda) \equiv M(\lambda+c) \equiv A_0(c)\lambda^m + \cdots + A_m(c). \qquad \text{(A.4)}$$

Furthermore, if $\rho \neq 0$ is a latent root of $M(\lambda)$, then $1/\rho$ is a latent root of

$$M^R(\lambda) \equiv \lambda^m M\left(\frac{1}{\lambda}\right) \equiv A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0. \quad \text{(A.5)}$$

If $M^R(\lambda)$ has a zero latent root (A_0 is singular), then $M(\lambda)$ is said to have an unbounded latent root. A lambda-matrix $M(\lambda)$ is said to be degenerate if $\det M(\lambda) = 0$ for all λ. This can only occur if A_0 and A_m are singular.

Consider the following algorithm for a non-degenerate lambda-matrix. It transforms a lambda-matrix with a singular leading matrix coefficient into one which is not. The transformed lambda-matrix is either

(i) $M^R(\lambda)$ if A_m is nonsingular

or (ii) $M^S(\lambda) \equiv \lambda^m M\left(\frac{1}{\lambda} + c\right)$, where $\det M(c) \neq 0$.

Part (ii) works since $\lambda^m M^S\left(\frac{1}{\lambda}\right) = M(\lambda+c)$, which does not have a zero latent root.

APPENDIX B

Previously Known Methods for Lambda-Matrices

The determinant of a lambda-matrix is a scalar polynomial. Let $f(\lambda) = \det M(\lambda)$. If one is willing to evaluate the determinant many times, then one can use any of a number of algorithms for the zeros of a scalar function. Tarnove [19] considers the use of Muller's method. He deflates known roots by considering $f_p(\lambda) = f(\lambda) \prod_{i=1}^{P-1} (\lambda-\lambda_i)^{-1}$.

Lancaster [10] notes that $f'(\lambda) = f(\lambda)\mathrm{Trace}\{M^{-1}(\lambda)M'(\lambda)\}$, which he uses in Newton's method. Newton's method is also used by Kublanovskaya [9], who finds $f(\lambda_i)/f'(\lambda_i)$ by using a factorization of $M(\lambda_i)$.

Another approach analyzed by Lancaster [12] is the use of a power-like method with a generalized Rayleigh quotient. That is, for arbitrary ξ_o, η_o and λ_o, let $\xi_i = [M(\lambda_i)]^{-1}\xi_o$, $\eta_i = \left[M^T(\lambda_i)\right]^{-1}\eta_o$, and $\lambda_{i+1} = \lambda_i - \dfrac{\eta_i^T M(\lambda_i)\xi_i}{\eta_i^T M'(\lambda_i)\xi_i}$. Lancaster has shown that, for a class of lambda-matrices, this iterative process is locally convergent and quadratic. Modifications of the above algorithm by $\xi_i = [M(\lambda_i)]^{-1}\xi_{i-1}$, $\eta_i = \left[M^T(\lambda_i)\right]^{-1}\eta_{i-1}$ has also been considered by Lancaster.

Another approach, due to Lancaster [14], is to consider the eigenvalues of $M(\lambda)$. Let $\mu(\lambda)$ be a scalar such that $M(\lambda) - \mu(\lambda)I$ is singular. Then a scalar ρ is needed such that $\mu(\rho) = 0$. Lancaster considers Newton's method on $\mu(\lambda)$.

The above methods of Lancaster and Kublanovskaya are only locally convergent and they do not have a method of deflation associated with them.

A symbol-manipulation approach is to perform Gaussian elimination on the lambda-matrix using polynomials in the computations. That is, every non-trivial lambda-matrix with $\det A_0 \neq 0$ can be transformed, by elementary transformations only, into a form such that $M(\lambda) = P(\lambda)N(\lambda)Q(\lambda)$, where $\det P(\lambda) = c_1 \neq 0$, $\det Q(\lambda) = c_2 \neq 0$ and $N(\lambda) = \mathrm{diag}\left(a_1(\lambda),\cdots,a_n(\lambda)\right)$, with $a_i(\lambda)$ monic polynomials and $a_i(\lambda)$ divides $a_{i+1}(\lambda)$. $N(\lambda)$ is called the Smith canonical form of $M(\lambda)$. See Wilkinson [22, p. 19]. Then all the roots of the $a_i(\lambda)$'s are latent roots of $M(\lambda)$.

This method parallels the approach of finding the characteristic equation in the eigenvalue problem.

APPENDIX C

The Quadratic Matrix Polynomial

The monic, quadratic matrix polynomial,

$$M(X) \equiv X^2 + A_1X + A_2, \tag{C.1}$$

with right solvents S_1 and S_2, is of the general form

$$M(X) = X^2 - \left[S_1 + (S_1-S_2)S_2(S_1-S_2)^{-1}\right]X + (S_1-S_2)S_2(S_1-S_2)^{-1}S_1 \tag{C.2}$$

if det $V(S_1,S_2)$ = det $(S_2-S_1) \neq 0$. Note that if S_1 and S_2 commute, then

$$M(X) = X^2 - (S_1+S_2)X + S_1S_2 \tag{C.3}$$

even if $V(S_1,S_2)$ is singular.

The corresponding lambda-matrix can be factored as

$$\begin{aligned} M(\lambda) &= \left(I\lambda - (S_1-S_2)S_2(S_1-S_2)^{-1}\right)(I\lambda-S_1) \\ &= \left(I\lambda - (S_1-S_2)S_1(S_1-S_2)^{-1}\right)(I\lambda-S_2). \end{aligned} \tag{C.4}$$

Thus,

$$R_2 = (S_1-S_2)S_2(S_1-S_2)^{-1} \tag{C.5}$$

and

$$R_1 = (S_1 - S_2) S_1 (S_1 - S_2)^{-1} \tag{C.6}$$

are left solvents of $M(X)$. From equation (C.5) it follows that

$$S_2^2 - S_1^2 = (S_1 + R_2)(S_2 - S_1). \tag{C.7}$$

Furthermore, $-A_1 = R_2 + S_1 = R_1 + S_2$ and $A_2 = R_2 S_1 = R_1 S_2$.

It is easily verified that

$$\begin{pmatrix} 0 & I \\ -A_2 & -A_1 \end{pmatrix} \begin{pmatrix} I & 0 \\ S_1 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ S_1 & I \end{pmatrix} \begin{pmatrix} S_1 & I \\ 0 & S_2 \end{pmatrix} \tag{C.8}$$

and hence, the block companion matrix is similar to $\begin{pmatrix} S_1 & I \\ 0 & S_2 \end{pmatrix}$ regardless of $V(S_1, S_2)$.

Assume that A_1 and A_2 are real matrices and let $S_1 = S_1^r + iS_1^c$ be a right solvent. Then,

$$M(\lambda) \equiv I\lambda^2 + A_1\lambda + A_2 = \left(I\lambda - \left(R_2^r + iR_2^c\right)\right)\left(I\lambda - \left(S_1^r + iS_1^c\right)\right). \tag{C.9}$$

Equating coefficients, we get $R_2^c + S_1^c = \underline{0}$ and $R_2^c S_1 + R_2^r S_1^c = \underline{0}$. Then, $R_2^r R_2^c + S_1^c S_1^r = \underline{0}$. By direct substitution it now follows that $S_1^r - iS_1^c$ is also a right solvent. Thus, Theorem C.1 *For a real, monic and quadratic matrix polynomial, if* $S + iC$ *is a right solvent, then*

(i) $S - iC$ *is a right solvent,*

(ii) $R - iC$ *is a left solvent, and*

(iii) $R + iC$ *is a left solvent,*

where $R + S = -A_1$.

Given arbitrary matrices S_1 and S_2, Corollary 4.1 shows that there might not be a monic, quadratic matrix polynomial having them as solvents. Such a condition occurs if S_1 and S_2 have distinct and disjoint eigenvalues and if $\det V(S_1,S_2) = 0$. If $V(S_1,S_2)$ is nonsingular, then $M(X)$ always exists. The following result gives necessary and sufficient conditions for the existence of $M(X)$.

Theorem C.2 *There exists a matrix polynomial* $M(X) = X^2 + A_1X + A_2$ *having right solvents* S_1 *and* S_2 *if and only if there exists a solution* Y *of*

$$Y(S_2-S_1) = \left(S_2^2-S_1^2\right). \tag{C.10}$$

Proof: In finding A_1 and A_2 to satisfy

$$M(S_1) = S_1^2 + A_1S_1 + A_2 = \underline{0}$$

$$M(S_2) = S_2^2 + A_1S_2 + A_2 = \underline{0} \tag{C.11}$$

the matrix A_1 must satisfy $A_1(S_2-S_1) = \left(S_2^2-S_1^2\right)$. #

Note that if $V(S_1,S_2)$ is singular and the condition of Theorem C.2 is satisfied, then $M(X)$ exists, but is not unique. From equation (C.10) if follows that

Corollary C.1 <u>If (S_2-S_1) is singular and $\left(S_2^2-S_1^2\right)$ is nonsingular, then there is no monic, quadratic matrix polynomials having S_1 and S_2 as right solvents.</u>

APPENDIX D

Computer Programs

The computer program that was used for Algorithm 1 follows. It is written in APL for the IBM 360/67. It is an interactive language and the program asks for

(i) the degree of the matrix polynomial,

(ii) the dimension of the matrix coefficients,

(iii) the matrix polynomial,

and (iv) the stopping criterion (an ε such that $\|M(X_1)\| < \varepsilon$ terminates the computation).

Following the code is an actual output for Example 1 in Chapter 9.

```
      ∇MAIN[⎕]∇
    ∇ MAIN;I;N;N1;X;X1;E1;E2;GITER;G;G1;M;N;M;SW
[1]    'INPUT DEGREE OF MATRIX POLYNOMIAL'
[2]    M←⎕
[3]    'INPUT SIZE OF MATRICES'
[4]    N←⎕
[5]    'MATRIX POLYNOMIAL'
[6]    M←((M+1),N,N)⍴⎕
[7]    ⎕←M←PAR M
[8]    GITER←0
[9]    E1←0.05
[10]   'ACCEPT ||M(S)||<'
[11]   E2←⎕
[12]   I←1
[13]  CALG:'STAGE ONE  ****************'
[14]   SW←0
[15]   G←E1 GRE M
[16]   N1←NORM M F X
[17]   'STAGE TWO  ****************   ||M(X)||= ';N1
[18]   →(I=1)/LP
[19]   →(N1<N)/LP
[20]   X←X1
[21]   N1←N
[22]   'USE OLD ITERATE TO START       ||M(X)||= ';N
[23]   X
[24]  LP:X1←(G F X)+.×INVP G1 F X
[25]   N←NORM M F X1
[26]   'ITERATION ';I;' ...............   ||M(X)||= ';N
[27]   I←I+1
[28]   X1
[29]   →(N<E2)/END
[30]   →(SW<2)/CONT
[31]   →(N<0.25×N1)/CONT
[32]   E1←E1×0.5
[33]   →(GITER<200)/CALG
[34]   'TOO MANY ITERATES'
[35]   →END
[36]  CONT:X←X1
[37]   N1←N
[38]   SW←SW+1
[39]   →LP
[40]  END:'******************************'
[41]   GITER+1;' ITERATIONS OF STAGE 1'
[42]   I-1;'  ITERATIONS OF STAGE 2'
[43]   'SOLVENT';X1
[44]   '||M(S)||= ';NORM M F X1
    ∇
```

```
      ∇GEE[⎕]∇
    ∇ G←E1 GEE M;SH;T;SW
[1]    SW←0
[2]    →(GITER>0)/LOOP
[3]    G1←(M,N,N)⍴(,(⍳N)∘.=⍳N),(0×⍳(M-1)×N×N)
[4]   LOOP:SH←(M,N,N)⍴(,G1[1+⍳(M-1);;]),(0×⍳N×N)
[5]    G←SH-G1[1;;] MULT M[1+⍳M;;]
[6]    →(G1[1;1;1]=0)/ZERO
[7]    'LATENT ROOT EST  ';GITER+1;'  ';G[1;1;1]÷G1[1;1;1]
[8]   ZERO:G←G÷NORM((M×N),N)⍴G
[9]    T←NORM((M×N),N)⍴G-G1
[10]   G1←G
[11]   SW←SW+1
[12]   GITER←GITER+1
[13]   →(GITER≥200)/TOOMUCH
[14]   →(SW>8)/TOOMUCH
[15]   →(E1<T)/LOOP
[16] TOOMUCH:SH←(M,N,N)⍴(,G1[1+⍳(M-1);;]),(0×⍳N×N)
[17]   G←SH-G1[1;;] MULT M[1+⍳M;;]
[18]   →(G1[1;1;1]=0)/ZERO1
[19]   'LATENT ROOT EST  ';GITER+1;'  ';G[1;1;1]÷G1[1;1;1]
[20] ZERO1:'FINISHED GENERATION OF G TO ';GITER+1
[21]   →(0=DETERM G1[1;;])/INIT
[22]   X←G[1;;]+.×INVP G1[1;;]
[23]   'INITIAL X ';X
[24]   →100
[25] INIT:X←(N,N)⍴?(0<⍳(N×N))×10
[26]   'ARTIFICIAL INITIAL X';X
[27]   →(0=DETERM G1 F X)/INIT
    ∇
```

```
      ∇F[⎕]∇
    ∇ A←M F S;I;J
[1]    I←1↑⍴M
[2]    A←M[1;;]
[3]    J←2
[4]   LP:A←M[J;;]+A+.×S
[5]    →(I≥J←J+1)/LP
    ∇
```

```
      ∇MULT[⎕]∇
    ∇ PROD←C MULT M;K;J
[1]    PROD←(J←⍴M)⍴0
[2]    K←1
[3]   LOOP:PROD[K;;]←C+.×M[K;;]
[4]    →(J[1]≥K←K+1)/LOOP
    ∇
```

```
      ∇PAR[⎕]∇
    ∇ MB←PAR M;J;K;T
[1]   MB←Jρ1,(0=ι2+J←ρM)
[2]   T←INVP M[1;;]
[3]   K←2
[4]  LOOP:MB[K;;]←T+.×M[K;;]
[5]   →(J[1]≥K←K+1)/LOOP
    ∇
```

```
      ∇NORM[⎕]∇
    ∇ N←NORM S
[1]   N←⌈/+/|S
    ∇
```

```
      ∇DETERM[⎕]∇
    ∇ D←DETERM M;J;K;I;T;N;U;L;V
[1]   U←M
[2]   L←(ιN)∘.=ιN←1↑ρM
[3]   V←ιN
[4]   J←1
[5]   D←1
[6]  NEXTCOL:V[J]←K←1↑⍒|U[J;]
[7]   V[K]←J
[8]   →(K=J)/NOCHANGE
[9]   D←-D
[10]  T←U[J;]
[11]  U[J;]←U[K;]
[12]  U[K;]←T
[13] NOCHANGE:I←J+1
[14]  →(1E¯10>|U[J;J])/SING
[15]  D←D×U[J;J]
[16]  →(J=N)/100
[17] NEXTROW:L[I;J]←U[I;J]÷U[J;J]
[18]  U[I;]←U[I;]-L[I;J]×U[J;]
[19]  →(N≥I←I+1)/NEXTROW
[20]  →(N>J←J+1)/NEXTCOL
[21]  →NOCHANGE
[22] SING:D←0
    ∇
```

```
      MAIN
INPUT DEGREE OF MATRIX POLYNOMIAL
⎕:    3
INPUT SIZE OF MATRICES
⎕:    2
MATRIX POLYNOMIAL
⎕:    1 0 0 1 ¯6 6 ¯3 ¯15 2 ¯42 21 65 18 66 ¯33 ¯81

  1   0
  0   1

 ¯6   6
 ¯3 ¯15

  2 ¯42
 21  65

 18  66
¯33 ¯81
ACCEPT ||M(S)||<
⎕:    .00001

STAGE ONE  ****************
LATENT ROOT EST  1  6
LATENT ROOT EST  2  2.666666667
LATENT ROOT EST  3  ¯3.75
LATENT ROOT EST  4  21.23333333
LATENT ROOT EST  5  9.791208791
FINISHED GENERATION OF G TO 5
INITIAL X
      3.992462543      ¯2.426142109
      1.213071054       7.631675705
STAGE TWO  ****************  ||M(X)||= 9.607873686
ITERATION 1 ...............  ||M(X)||= 1.277362968
      3.972923527      ¯2.089215678
      1.044607839       7.106747044
ITERATION 2 ...............  ||M(X)||= 0.2293107142
      3.992690243      ¯2.017863253
      1.008931626       7.019485122
```

```
ITERATION 3 ...............  ||M(X)||= 0.04157477338
      3.99847172        ¯2.00035881
      1.001679405        7.003509934
ITERATION 4 ...............  ||M(X)||= 0.007424137007
      3.999709525       ¯2.000600132
      1.000304566        7.000623223
ITERATION 5 ...............  ||M(X)||= 0.001311183582
      3.999947122       ¯2.000108385
      1.000054192        7.000109699
ITERATION 6 ...............  ||M(X)||= 0.0002301117362
      3.999990575       ¯2.000019094
      1.000009547        7.000019217
ITERATION 7 ...............  ||M(X)||= 4.024636788E¯5
      3.999998338       ¯2.000003346
      1.000001673        7.000003358
ITERATION 8 ...............  ||M(X)||= 7.026157959E¯6
      3.999999709       ¯2.000000585
      1.000000292        7.000000586
****************************
5 ITERATIONS OF STAGE 1
8  ITERATIONS OF STAGE 2
SOLVENT
      3.999999709       ¯2.000000585
      1.000000292        7.000000586
||M(S)||= 7.026157959E¯6
```

REFERENCES

1. P. Franklin, "Algebraic matric equations," J. Math. Physics, 10(1931), 289-314.

2. F. Gantmacher, The Theory of Matrices. I, II, Chelsea Publishing Co., New York, 1960.

3. K. Guderley, "On nonlinear eigenvalue problems for matrices," SIAM J. Appl. Math., 6(1958), 335-353.

4. A. Householder, "Generalizations of an algorithm of Sebastião e Silva," Numerische Mathematik, 16(1971), 375-382.

5. A. Householder, The Theory of Matrices in Numerical Analysis. Blaisdell, New York, 1964.

6. A. Jameson, "Solution of the equation AX + XB = C by inversion of an N×N or N×N matrix," SIAM J. Appl. Math., 16(1968), 1020-1023.

7. M. Jenkins and J. Traub, "A three-stage algorithm for real polynomials using quadratic iteration," SIAM J. Numer. Anal., 7(1970), 545-566.

8. M. Jenkins and J. Traub, "A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration," Numerische Mathematik, 14(1970), 252-261.

9. V. Kublanovskaya, "On an approach to the solution of the generalized latent value problem for λ-matrices," SIAM J. Numer. Anal., 7(1970), 532-537.

10. P. Lancaster, "Algorithms for lambda-matrices," Numerische Mathematik, 6(1964), 388-394.

11. P. Lancaster, "Explicit solutions of linear matrix equations," SIAM Review, 12(1970), 544-566.

12. P. Lancaster, "A generalised Rayleigh quotient iteration for lambda-matrices," Arch. Rat. Mech. & Anal., 8(1961), 309-322.

13. P. Lancaster, Lambda-matrices and Vibrating Systems. Pergamon Press, New York, 1966.

14. P. Lancaster, "Some applications of the Newton-Raphson method to non-linear matrix problems," Proc. Royal Soc. London, (A), 271(1963), 324-331.

15. C. MacDuffee, The Theory of Matrices. Chelsea, New 1946.

16. J. Ortega and W. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970.

17. G. Peters and J. Wilkinson, "AX = ABX and the gene alized eigenproblem," SIAM J. Num. Anal. 7(19 479-492.

18. W. Roth, "On the unilateral equation in matrices," Trans. Amer. Math. Soc., 32(1930), 61-80.

19. I. Tarnove, "Determination of eigenvalues of matri having polynomial elements," J. Soc. Indust. Math., 6(1958), 163-171.

20. J. Traub, "The calculation of zeros of polynomials analytic functions," Mathematical Aspects of Computer Science, Proc. Symp. Appl. Math., 19 American Mathematical Society, Providence, R. 1967, 138-152.

21. J. Traub, "A class of globally convergent iteratic functions for the solution of polynomial equa Mathematics of Computation, 20(1966), 113-13

22. J. Wilkerson, The Algebraic Eigenvalue Problem. C London, 1965.

York,

r-
70),

ces
Appl.

and

,
I.,

n
tions,"

anford,