GAUSS ELIMINATION OPERATOR FOR DETERMINING THE CHARACTERISTIC POLYNOMIAL AND EIGENVECTORS OF ANY SQUARE MATRIX

by

CHI-CHING CHEN

B.S., National Tsing Hua University, 1969

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1972

Approved by:

R. F. Hincovec

Major Professor

TABLE OF CONTENTS

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to the following persons: Dr. R. F. Sincovec, major professor, for his constructive guidance, precious criticism and deep interest in preparing this thesis, Dr. R. H. Hetherington for his valuable comments, and Dr. Shih-Lu Chang for his enlightened suggestions. The author dedicates this thesis to his parents and his fiancee.

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

Chapter 1

INTRODUCTION

1. What is the problem?

Let us denote the characteristic polynomial of a n x n matrix A by

$$(-1)^n \; [\lambda^n - P_1\lambda^{n-1} - P_2\lambda^{n-2} \dots - P_n] = 0.$$

Then it can be easily verified that the characteristic polynomial of the matrix

$$L_1 = \left\| \begin{matrix} P_1 & P_2 & \dots & P_{n-1} & P_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 1 & 0 \end{matrix} \right\|$$

or

$$L_2 = \left\| \begin{matrix} 0 & 0 & \dots & 0 & P_n \\ 1 & 0 & \dots & 0 & P_{n-1} \\ 0 & 1 & \dots & 0 & P_{n-2} \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 1 & P_1 \end{matrix} \right\|$$

is identical to the characteristic polynomial of A. The matrix L_1 (or L_2) is called the companion matrix of the characteristic polynomial of A.

It is very natural to ask whether L_1 (or L_2) is similar to A. Suppose the matrix A has distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then we know that A is similar to diag (λ_i) and it can be shown that in this case diag (λ_i) is similar to the companion matrix. (The proof is omitted). Suppose there are some multiple eigenvalues and that A has more than one independent eigenvector corresponding to λ_i, then it can not be similar to L_1. For a matrix to be similar to the companion matrix of its characteristic polynomial, it is necessary that its Jordan canonical form should have only one Jordan submatrix associated with each distinct eigenvalue.

A matrix is called "derogatory" if there is more than one Jordan submatrix (and therefore more than one eigenvector) associated with λ_i for some i. Such a derogatory matrix cannot be similar to the companion matrix of its characteristic polynomial, and we now investigate the canonical form analogous to the companion form for a derogatory matrix. In other words, our problem is to identify the canonical form analogous to the companion form for a general matrix (derogatory or not derogatory), and to develop the similarity transformation between A and its "analogous" companion matrix. Such a similarity transformation can be derived from the Gauss elimination operator.

2. Review of Literature.

By defining a Frobenius matrix B_r of order r to be a matrix of the form

$$B_r = \left\|\begin{matrix} b_1 & b_2 & \cdots & b_{r-1} & b_r \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & & 1 & 0 \end{matrix}\right\|,$$

the fundamental theorem for general matrices may be expressed in the following form:

Every matrix A may be reduced by a similarity transformation to the direct sum of a number S of Frobenius matrices which may be denoted by $B_{r_1}, B_{r_2}, \ldots B_{r_S}$ where the characteristic polynomial of each B_{r_i} divides the characteristic polynomial of all the preceding B_{r_j}. In the case of a non-derogatory matrix $S = 1$ and $r_1 = n$.

The proof for the existence of the Frobenius canonical form is omitted. The important thing we wish to discuss is how we can reduce the matrix A to the Frobenius canonical form* and what kind of similarity transformation is most convenient for computing.

A two-stage reduction of a general matrix A to Frobenius form has been described by J.H. Wilkinson [Ref. 1].

Stage 1. A is reduced to upper Hessenberg form H by stabilized elementary transformations.

Stage 2. H is reduced to Frobenius form by non-stabilized elementary transformations.

---

*The direct sum of the Frobenius matrices is called the Frobenius canonical form.

Danilewski [Ref. 2] combined these two stages to form a method consisting of n-1 steps.

3. What is the new approach?

In this thesis, a different approach is being developed. We propose a canonical form analogous to the companion form for a general matrix that consists not only of a direct sum of a number S of Frobenius matrices but also some "additional" coefficients in some corresponding rows. That is L*, as follows

$$L^* = \left\| \begin{matrix} B_{r_1} & C_{12} & C_{13} & C_{14} & \cdots & C_{1s} \\ & B_{r_2} & C_{23} & C_{24} & \cdots & C_{2s} \\ & & D_{r_3} & F_{34} & \cdots & F_{3s} \\ & & & B_{r_4} & \cdots & C_{4s} \\ & 0 & & & \vdots & \vdots \\ & & & & & C_{s-1\,s} \\ & & & & & B_{r_s} \end{matrix} \right\|$$

where B_{r_i} is a Frobenius matrix with order r_i,

C_{ij} is the zero matrix except for the first row which may contain nonzero entries,

D_{r_i} is an upper tridiagonal matrix,

F_{ij} is an arbitrary matrix.

It is a more generalized case than the Frobenius canonical form in the previous fundamental theorem which may be a special case when all the C_{ij}, D_{r_i} and F_{ij} are all zero and the characteristic polynomial of each B_{r_i} divides the characteristic polynomial of all the preceding B_{r_j}

We claim the fundamental theorem for general matrices may be expressed in the following form.

Theorem 1.1: Every matrix A may be reduced by a similarity transformation to the matrix L* (L* is the matrix stated above). In addition, such a similarity transformation can be derived from "modified" Gauss elimination operator.

The proof for the existence of the fundamental theorem and the development of such a similarity transformation will be described in the next chapter.

Chapter 2

ALGORITHMS

1. Development of a special similarity transformation.

It has been mentioned in Chapter 1 that a similarity transformation can be derived from the "modified" Gauss elimination operator.

Let us review some interesting properties of Gauss elimination. Define

(1) $A \equiv (a_{ij})$ to be a given matrix and

(2) $A^{(K)} \equiv (a_{ij}^{(K)})$.

For $K = 1$ we have $A^{(1)} \equiv A$, and the elements in (2) for $K = 2,3, \ldots,n$ are computed recursively by

$$(3)\quad a_{ij}^{(K)} = \begin{cases} a_{ij}^{(K-1)} & \text{for } i \le K-1, \\ 0 & \text{for } i \ge K,\ j \le K-1, \\ a_{ij}^{(K-1)} - \dfrac{a_{i,K-1}^{(K-1)}}{a_{K-1,K-1}^{(K-1)}}\, a_{K-1,j}^{(K-1)} & \text{for } i \ge K,\ j \ge K. \end{cases}$$

The resulting coefficient matrix has the form:

$$
A^{(K)} \equiv \left\| \begin{array}{cccccccc}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1,K-1}^{(1)} & a_{1K}^{(1)} & \cdots & a_{1n}^{(1)} \\
0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2,K-1}^{(2)} & a_{2K}^{(2)} & \cdots & a_{2n}^{(2)} \\
 & & & & \vdots & \vdots & & \vdots \\
 & & & & a_{K-1,K-1}^{(K-1)*} & a_{K-1,K}^{(K-1)} & & a_{K-1,n}^{(K-1)} \\
 & 0 & & & 0 & a_{KK}^{(K)} & \cdots & a_{Kn}^{(K)} \\
 & & & & \vdots & \vdots & & \vdots \\
 & & & & 0 & a_{nK}^{(K)} & \cdots & a_{nn}^{(K)}
\end{array} \right\| .
$$

In terms of an operator to reduce $A^{(K-1)}$ to $A^{(K)}$, it is

$$
(4) \qquad
\begin{array}{c}
\begin{array}{ccccccc}
 & & \text{K-1th column} & & & & \\
\end{array} \\
\left\| \begin{array}{ccccccc}
1 & & 0 & & & & \\
 & \ddots & \vdots & & & & \\
 & 1 & 0 & & & & \\
 & & \vdots & & & & \\
 & & 1 & & & & \\
 & & -a_{K,K-1}^{(K-1)}/a_{K-1,K-1}^{(K-1)} & 1 & & & \\
 & & -a_{K+1,K-1}^{(K-1)}/a_{K-1,K-1}^{(K-1)} & & 1 & & \\
 & & \vdots & & & \ddots & \\
 & & -a_{n,K-1}^{(K-1)}/a_{K-1,K-1}^{(K-1)} & & & & 1
\end{array} \right\|
\end{array}
A^{(K-1)} \equiv A^{(K)} .
$$

Usually the matrix on the left side of equation (4) is called the Gauss elimination operator. (Note: Because the coefficients of this matrix depend on the matrix $A^{(K-1)}$, we call it an "operator" instead of a "transformation").

---

* $a_{K-1,K-1}^{(K-1)}$ is the pivot element when we reduce $A^{(K-1)}$ to $A^{(K)}$; this element can not be zero.

Along the elimination idea, the Gauss elimination operator can be extended in three ways:

1. The (K-1)th column can be defined to eliminate above the diagonal element as well as below.

2. The pivot element is not necessarily restricted to the diagonal element which is $a_{K-1,K-1}^{(K-1)}$ in the above matrix, but can be any element in the (K-1)th column of $A^{(K-1)}$.

3. The modified (K-1)th column can replace any column in the identity matrix.

Let us call the modified matrix a modified Gauss elimination operator on an arbitrary matrix A which is given by

$$E_{K,A}^{j} = \left\| \begin{array}{cccccc} 1 & & & -a_{1j}/a_{Kj} & & \\ & 1 & & -a_{2j}/a_{Kj} & & \\ & & 1 & \cdots & & \\ & & & -a_{jj}/a_{Kj} & & \\ & & & \cdots & & \\ & & & 1/a_{Kj} & & \\ & & & \cdot\cdot & \cdot & \\ & & & -a_{nj}/a_{Kj} & & 1 \end{array} \right\| \quad \text{provided } a_{Kj} \neq 0.$$

K-th column

Note: The nonzero column is derived from the j-th column of A, but is put in the K-th column of the elimination operator.

Note: Subscript A in $E_{K,A}^{j}$ will be omitted if no confusion occurs.

E_K^j can also be defined as

$$E_K^j = I_n + \frac{1}{a_{Kj}} (e_K\, e_j^T - A)\; e_j\; e_K^T$$

where T denotes transpose,

$$e_K^T = \|\; 0,\; 0,\; \ldots\; 0,\; \underset{K\text{-th}}{1},\; 0,\; \ldots\; 0 \;\|, \text{ and}$$

$$1 \le j,\; K \le n.$$

Several properties of the operator E_K^j will be examined in the following discussion.

Theorem 2.1: The inverse of E_K^j, denoted by $(E_K^j)^{-1}$, exists and is equal to

$$I_n + (A - e_K\, e_j^T)\; e_j\; e_K^T, \qquad 1 \le j,\; K \le n\;.$$

Proof: $$(E_K^j)^{-1}(E_K^j) = (I_n + (A - e_K\, e_j^T)\; e_j\; e_K^T)(I_n + \frac{1}{a_{Kj}} (e_K\, e_j^T - A) e_j\; e_K^T)$$

$$= I_n + (A - e_K\, e_j^T)\; e_j\; e_K^T + \frac{1}{a_{Kj}} (e_K e_j^T - A)\; e_j\; e_K^T$$

$$+ \frac{1}{a_{Kj}} (A - e_K\, e_j^T)\; e_j\; e_K^T\; (e_K\, e_j^T - A)\; e_j\; e_K^T$$

$$= I_n + a_j\, e_K^T - e_K\, e_K^T + \frac{1}{a_{Kj}}\, e_K\, e_K^T - \frac{1}{a_{Kj}}\, a_j\, e_K^T$$

$$+ \frac{1}{a_{Kj}} (a_j\, e_K^T - e_K\, e_K^T + a_j (e_K^T\, a_j)\; e_K^T$$

$$+ e_K (e_K^T\, a_j)\; e_K^T)$$

$$= I_n.$$

Note: a_i means the ith column of matrix A and

a_{Kj} means the element in the (K,j) position of matrix A.

Prop. 1: $E_K^j\ e_i = e_i$ when $i \neq K$

$$= e_K + \frac{1}{a_{Kj}} (e_K - a_j) \quad \text{when } i = K.$$

Proof: $E_K^j\ e_i = (I_n + \frac{1}{a_{Kj}} (e_K\ e_j^T - A)\ e_j\ e_K^T)\ e_i$

$$= I_n\ e_i + \frac{1}{a_{Kj}} (e_K\ e_j^T - A)\ e_j\ \delta_{Ki}$$

$$= \begin{cases} I_n\ e_i = e_i & \text{for } i \neq K \\ I_n\ e_K + \frac{1}{a_{Kj}} (e_K - a_j) & \text{for } i = K. \end{cases}$$

Prop. 2: $E_K^j\ a_j = e_K$

Proof: $E_K^j\ a_j = [I_n + \frac{1}{a_{Kj}} (e_K\ e_j^T - A)\ e_j\ e_K^T]\ a_j$

$$= a_j + \frac{1}{a_{Kj}} (e_K\ e_j^T - A)\ e_j\ a_{Kj}$$

$$= a_j + e_K(e_j^T\ e_j) - a_j$$

$$= e_K.$$

Note: Prop.2 shows that E_K^j will reduce the j-th column of A to e_K.

Prop.3: $A(E_K^j)^{-1} = (a_1, a_2, \ldots, a_{K-1}, (A)^2\ e_j, a_{K+1}, \ldots, a_n).$

Proof: $A(E_K^j)^{-1} = A(I_n + (A - e_K\ e_j^T)\ e_j\ e_K^T)$

$$= A + A(Ae_j)\, e_K^T - Ae_K\, (e_j^T\, e_j)\, e_K^T$$

$$= A + (A^2)\, e_j\, e_K^T - a_K\, e_K^T$$

$$= (a_1,\, a_2,\, \ldots,\, a_{K-1},\, (A^2)\, e_j,\, a_{K+1},\, \ldots,\, a_n).$$

Note: Prop. 3 shows that if A is multiplied by $(E_K^j)^{-1}$ on the right, it replaces the K-th column of A by the j-th column of A^2 and does not change the other columns.

Theorem 2.2: If $K \neq j$ and $a_{Kj} \neq 0$ then

$$E^j_{K,A*}\, [A(E^j_{K,A})^{-1}] = [E^j_{K,A}\, A]\, (E^j_{K,A^+})^{-1}$$

where $A* = A(E^j_{K,A})^{-1}$ and $A^+ = E^j_{K,A}\, A$

$E^j_{K,A*}$ and E^j_{K,A^+} are modified Gauss elimination operator defined on $A*$ and A^+ respectively.

Before proving Theorem 2.2, we prove the following Lemma 2.1.

Lemma 2.1: If $K \neq j$ and $a_{Kj} \neq 0$ then

$$E^j_{K,A*} = E^j_{K,A} \text{ and } (E^j_{K,A^+})^{-1} = (E^j_{K,A})^{-1}.$$

Proof: Since $A* = A\, (E^j_{K,A})^{-1}$ and $A^+ = E^j_{K,A}\, A$ it follows that

$$(A*)_j = (A)_j = a_j \text{ by Prop. 3.}$$

$$(A^+)_j = (E^j_{K,A}\, A)_j = e_j$$

Then by the definition of E_K^j and $(E_K^j)^{-1}$ we have

$$E^j_{K,A*} = E^j_{K,A} \text{ and } (E^j_{K,A}+)^{-1} = (E^j_{K,A})^{-1}.$$

Proof of Theorem 2.2: Lemma 2.1 immediately shows us that

$$E^j_{K,A*} [A(E^j_{K,A})^{-1}]$$

$$= E^j_{K,A} [A(E^j_{K,A})^{-1}] = [E^j_{K,A} A] (E^j_{K,A})^{-1} = [E^j_{K,A} A](E^j_{K,A}+)^{-1}$$

Note: It follows that E^j_K (subscript A is omitted) can be defined as a similarity transformation with inverse matrix $(E^j_K)^{-1}$ when $K \neq j$ and $a_{Kj} \neq 0$.

Corollary 2.2: $(E^j_K a_K)(E^j_K)^{-1} = E^j_K A^2 e_j$

Proof:

$$E^j_{K,A*} [A(E^j_{K,A})^{-1}]$$

[5] $$= E^j_K [A(E^j_K)^{-1}]$$

$$= E^j_K (a_1, a_2, a_3, \ldots, a_j, \ldots, (A)^2 e_j, a_{K+1}, \ldots, a_n)$$

$$= (E^j_K a_1, \ldots, E^j_K a_{j-1}, e_K, E^j_K a_{j+1}, \ldots, E^j_K A^2 e_j, E^j_K a_{K+1} \ldots).$$

$$[E^j_{K,A} A] (E^j_{K,A}+)^{-1}$$

[6] $$= [E^j_{K,A} A](E^j_{K,A})^{-1}$$

$$= (E^j_K a_1, \ldots, E^j_K a_{j-1}, e_K, E^j_K a_{j+1}, \ldots$$

$$E^j_K a_{K-1}, (E^j_K a_K)(E^j_K)^{-1}, E^j_K a_{K+1}, \ldots, E^j_K a_n)$$

It follows that

$$(E_K^j\ a_K)(E_K^j)^{-1} = E_K^j(a_K(E_K^j)^{-1})$$

$$= E_K^j\ (A)^2\ e_j$$

since Theorem 2.2 implies that (6) is identical to (5)

Note: The proof of Corollary 2.2 also shows us that the j-th column of A is replaced by e_K and $a_K(E_K^j)^{-1} = (A)^2\ e_j$.

Until now it was natural to consider E_K^j as a transformation instead of an operator. We now use E_K^j to define a similarity transformation which we will use to prove our fundamental theorem and to get a method for solving pratical problems.

2. "Analogous" companion matrix reducing process.

Some more properties are necessary before the reducing process is introduced. Whenever we use E_K^j in the following discussion, we assume that $K \neq j$ and $a_{Kj} \neq 0$ hold.

Lemma 2.2: If the j-th column of matrix A is e_K, then the j-th column of matrix

$$E_r^s\ A(E_r^s)^{-1}$$

is also e_K whenever $r \neq K$ and $r \neq s$.

Proof: $E_r^s\ A(E_r^s)^{-1}$

$$= E_r^s\ (a_1,\ a_2,\ \ldots,\ a_{j-1},\ e_K,\ a_{j+1},\ \ldots,\ a_n)(E_r^s)^{-1}$$

$$= E_r^s\,(a_1, a_2, \ldots, a_{j-1}, e_K, a_{j+1}, \ldots, a_{r-1}, (A^2)\,e_s, a_{r+1}, \ldots, a_n)$$

$$= (E_r^s\,a_1, E_r^s\,a_2, \ldots, E_r^s\,a_{j-1}, e_K, E_r^s\,a_{j+1}, \ldots,$$

$$E_r^s\,a_{s-1}, e_r, E_r^s\,a_{s+1}, \ldots, E_r^s\,(A^2)\,e_s, E_r^s\,a_{r+1}, \ldots, E_r^s\,a_n).$$

Note: Lemma 2.2 shows us that if we define such a similarity transformation twice with $r \neq K$, $r \neq s$ and $K \neq j$, then A will be reduced to a matrix with two unit vectors, e_K in the j-th column and e_s in the r-th column.

Theorem 2.3: For any square matrix A, there exists a sequence of successive similarity transformations which are defined using the modified Gauss elimination operator and reduce A to L** where L** has the following form:

$$L^{**} = \left\|\begin{matrix} L_2^{r_1} & H_{12} & F_{13} & H_{14} & \cdots & H_{1s} \\ & L_2^{r_2} & F_{23} & H_{24} & \cdots & H_{2s} \\ & & D_{r_3} & H_{34} & \cdots & H_{3s} \\ & & & L_2^{r_4} & \cdots & H_{4s} \\ & 0 & & & \cdots & . \\ & & & & \cdots & . \\ & & & & & L_2^{r_s} \end{matrix}\right\|$$

where $L_2^{r_i}$ is a L_2 matrix with order r_i,

H_{ij} is a zero matrix except for the last column which may contain nonzero entries,

D_{r_i} is an upper tridiagonal matrix,

F_{ij} is an arbitrary matrix.

Note: L** and L* have the same characteristic polynomial. When L** just has the first block that means L** = L.

Proof: The sequence order of pivots is (2,1), (3,2), ..., (K+1,K) ... (n,n-1) where we define $E^K_{K+1,A^{K-1}}$ on A^{K-1}. Here A^o denotes the original A.

Case 1: If $(A^{K-1})_{K+1,K} \neq 0$ for K = 1, 2, 3, ..., n-1

then $E^1_{2,A^o} A^o (E^1_{2,A^o})^{-1} = (e_2, a^1_2, a^1_3 \dots a^1_n) = A^1$ at the first step, and

$$E^K_{K+1,A^{K-1}} A^{K-1} (E^K_{K+1,A^{K-1}})^{-1} = (e_2, \dots, e_{K+1}, a^K_{K+1}, \dots, a^K_n)$$

$$= A^K$$ at the K-th step. For

K = n - 1, the last step, we have

$$E^{n-1}_{n,A^{n-2}} A^{n-2} (E^{n-1}_{n,A^{n-2}})^{-1} = (e_2, e_3, \dots, e_n, a^{n-1}_n).$$

That is, we reach L** with L** = L.

Case 2: If there exists some zero pivot element which occurs after the K-th reduction operation, that is

$$(A^{K-1})_{K+1,K} = 0.$$

Then we check the rest of the coefficients in the K-th column to see if

$$(A^{K-1})_{K+p,K} = 0, \qquad p = 2,3, \dots, n-k.$$

Case 2.1: If there exists some p' such that

$$(A^{K-1})_{K+p',K} \neq 0$$

then we move this nonzero element to the (K+1,K) position by applying an elementary transformation which consists of a permutation matrix. The rest of the proof is the same as in Case 1.

Case 2.2: But if

$$(A^{K-1})_{K+p',\,K} = 0, \quad p' = 2,3,\ \ldots,\ n-K,$$

then we skip to search the next column, i.e. replace K by K' = K+1 and see if the elements of the matrix

$$(A^{K'-1})_{K'+p,\ K'} \neq 0 \quad \text{for some } p = 1,2,\ \ldots,\ n-K.$$

If so, go to Case 2.1; else, repeat Case 2.2 and terminate searching at $(A^{n-2})_{n,n-1}$. Finally we will get L**.

There is another way to define a modified Gauss elimination operator by the rows of any square matrix A. Let us call this method the second modified Gauss elimination operator given by

$$R^{j}_{K,A} = \begin{array}{c} \begin{array}{ccccc} \phantom{\frac{-a_{K1}}{a_{Kj}}} & \phantom{\frac{-a_{K2}}{a_{Kj}}} & & \text{jth column} & \phantom{\frac{-a_{Kn}}{a_{Kj}}} \end{array} \\ \left\| \begin{array}{ccccc} 1 & 0 & & & 0 \\ 0 & 1 & & & \\ & & 1 & & \\ & & & \cdot & \\ \frac{-a_{K1}}{a_{Kj}} & \frac{-a_{K2}}{a_{Kj}} & & \frac{1}{a_{Kj}} & \frac{-a_{Kn}}{a_{Kj}} \\ & & & & \cdot \\ & 0 & & & 1 \end{array} \right\| \end{array}.$$

Then, the inverse of $R^j_{K,A}$ is

$$(R^j_{K,A})^{-1} = \left\|\begin{matrix} 1 & 0 & & & & \\ 0 & 1 & & & 0 & \\ & & \cdot & & & \\ a_{K1} & a_{K2} & & a_{KK} & \cdots & a_{Kn} \\ & & & & 1 & \\ & & & & & \cdot \\ & 0 & & & & 1 \end{matrix}\right\|$$

and we can define a similarity transformation by the same approach as before, that is

$$(R^j_K)^{-1} AR^j_K, \qquad \text{if } K \neq j \quad \text{and } a_{Kj} \neq 0$$

and get a result like Theorem 2.2.

Theorem 2.3': For any square matrix A, there exists a sequence of successive similarity transformations which is defined using the second modified Gauss elimination operator and will reduce A to L*.

Proof: The proof is quite similar except it considers the sequence of pivots in the order (n, n-1), (n-1, n-2), ..., (2,1), for defining $R^{K-1}_{K,A^{(K-2)}}$. We skip the detailed proof.

Note: Theorem 1.1 will follow from Theorem 2.3.

3. Computing the eigenvector

We have now shown that for any matrix A and its "analogous companion matrix", there exists a similarity transformation which can be developed using the modified Gauss elimination operator or the second modified Gauss elimination operator.

Let us define the following notation:

$$E^K = E^K_{K+1,SA} \quad \text{if } (A)_{K+1,K} = 0 \quad \text{but } (SA)_{K+1,K} \neq 0,$$

(S is an elementary similarity transformation which satisfies Case 2.1 in Theorem 2.2).

$$= I_n \quad \text{if } (A)_{K+p,K} = 0 \quad \text{for all } p \geq 1,$$

$$= E^K_{K+1,A} \quad \text{otherwise.}$$

Then our process can be written

as $[E^{n-1} [\ldots [E^2 [E^1 A(E^1)^{-1}] (E^2)^{-1}] \ldots] (E^{n-1})^{-1}]$

$= L^{**}$.

Let $\pi E = (E^{n-1} \ldots E^2 E^1)$; then

it follows that

(7) $$(\pi E) A(\pi E)^{-1} = L^{**}. \qquad (3.1)$$

Theorem 2.4: Let $X_{L^{**},R}$ and $X_{L^{**},L}$ be a right eigenvector and a left eigenvector of L^{**}. Let $X_{A,R}$ and $X_{A,L}$ be a right eigenvector and a left eigenvector of A corresponding to the eigenvalues λ and ρ, respectively. Then i) $X_{A,R} = (\pi E)^{-1} X_{L^{**},R}$ and

ii) $X_{A,L} = X_{L^{**},L}(\pi E)$.

Proof: From (7),

(8) $$(\pi E)A = L^{**}\ (\pi E)$$ and it follows that

(9) $$A(\pi E)^{-1} = (\pi E)^{-1}\ L^{**}.$$

By (8), $(\pi E)\ AX_{A,R} = L^{**}\ (\pi E)\ X_{A,R}$ which implies

$$\lambda(\pi E)\ X_{A,R} = L^{**}\ (\pi E)\ X_{A,R}.$$

Hence $(\pi E)\ X_{A,R}$ is a right eigenvector of L^{**} corresponding to the eigenvalue λ. Also, since

$$X_{L^{**},R} = (\pi E)\ X_{A,R},$$

we have

$X_{A,R} = (\pi E)^{-1}\ X_{L^{**},R}$ which can be used in computing the eigenvectors of A from those of L^{**}.

By (9), $X_{A,L}\ A(\pi E)^{-1} = X_{A,L}(\pi E)^{-1}\ L^{**}$ which implies

(10) $$\rho X_{A,L}\ (\pi E)^{-1} = X_{A,L}\ (\pi E)^{-1}\ L^{**}.$$

Hence $X_{A,L}(\pi E)^{-1}$ is a left eigenvector of L^{**} corresponding to the eigenvalue ρ. From (10),

$X_{L^{**},L} = X_{A,L}\ (\pi E)^{-1}$ and therefore

$X_{A,L} = X_{L^{**},L}(\pi E)$ which concludes the proof of the theorem.

If we define R^K, in the same way as E^K, then we will get similar results, such as

$$L^* = [(R^1)^{-1}\ [\dots\ [(R^{n-2})^{-1}[(R^{n-1})^{-1}\ AR^{n-1}]\ R^{n-2}]\ \dots\ ]\ R^1].$$

Let $\pi R = (R^{n-1}\ R^{n-2}\ \dots\ R^2R^1)$ and

it follows that $(\pi R)^{-1}\ A(\pi R) = L^*$

and we still get a theorem analogous to Theorem 2.4.

Theorem 2.5: Let $X_{L*,R}$ and $X_{L*,L}$ be a right eigenvector and a left eigenvector of L*. Let $X_{A,R}$ and $X_{A,L}$ be a right eigenvector and a left eigenvector of A corresponding to eigenvalues λ and ρ:

i) $X_{A,R} = (\pi R)\, X_{L*,R}$ and

ii) $X_{A,L} = X_{L*,L}\, (\pi R)^{-1}$.

Proof: Omitted.

Since $X_{L*,R}$, $X_{L*,L}$, $X_{L**,L}$, and $X_{L**,R}$ can be easily calculated, and πR, $(\pi R)^{-1}$, πE and $(\pi E)^{-1}$ have been calculated while reducing A into L* or L**, we can then calculate $X_{A,R}$ and $X_{A,L}$ by Theorem 2.4 and Theorem 2.5.

Let us investigate the eigenvectors (right or left) of L* and L** [+] corresponding to any given eigenvalue. We restrict L** to the following form, but the more general case can be considered by only minor modifications in the following discussion:

[+] From now on, L* and L** are restricted to those which are derived from applying non-trivial reducing process to A, ie. $\pi E \neq I$ or $\pi E \neq S$, $\pi R \neq I$ or $\pi R \neq S$. See Chapter 4.

$$\left\|\begin{array}{cccc|cccc|cccc}
0 & 0 & & -\alpha_p & & & & -\beta_{m+p} & & & & -\gamma_{n+m+p} \\
1 & 0 & & -\alpha_{p-1} & & & & \vdots & & & & \vdots \\
0 & 1 & & -\alpha_{p-2} & & & & & & & & \\
 & \ddots & 1 & -\alpha_1 & & & & -\beta_{m+1} & & & & \\
 & & & & 0 & 0 & & -\beta_m & & & & -\gamma_{n+m} \\
 & & & & 1 & 0 & & -\beta_{m-1} & & & & \vdots \\
 & & & & 0 & 1 & & -\beta_{m-2} & & & & \\
 & & & & & \ddots & 1 & -\beta_1 & & & & -\gamma_{n+1} \\
 & & & & & & & & 0 & 0 & & -\gamma_n \\
 & & & & & & & & 1 & 0 & & \vdots \\
 & & & & & & & & 0 & 1 & & \\
 & & & & & & & & & \ddots & & -\gamma_2 \\
 & & & & & & & & & & 1 & -\gamma_1
\end{array}\right\|
\left\|\begin{array}{c} X_1 \\ \vdots \\ \\ X_p \\ Y_1 \\ \vdots \\ \\ Y_m \\ Z_1 \\ \vdots \\ \\ Z_n \\ \\ \end{array}\right\|
= \lambda
\left\|\begin{array}{c} X_1 \\ \vdots \\ \\ X_p \\ Y_1 \\ \vdots \\ \\ Y_m \\ Z_1 \\ \vdots \\ \\ Z_n \\ \\ \end{array}\right\|$$

The preceding equations lead directly to the following set of formula for calculating $X_{L**,R}$.

Method α: Set $Z_1 = 1$, $Z_n = -\lambda/\gamma_n$, and

$$Z_{n-K} = \gamma_K Z_n + \lambda Z_{n-K+1}, \qquad K = 1, \ldots, n-2.$$

Next, calculate Y_i.

Set $Y_1 = 1$, $Y_m = -(\lambda + \gamma_{n+m} Z_n)/\beta_m$,

$$Y_{m-K} = \beta_K Y_m + \gamma_{n+K} Z_n + \lambda Y_{m-K+1},$$

$$K = 1, 2, \ldots, m-2.$$

Next, calculate X_i.

Set $X_1 = 1, \qquad X_p = -(\lambda + \beta_{m+p}\, Y_m + Y_{m+n+p}\, Z_n)/\alpha_p,$

$$X_{p-K} = \alpha_K X_p + \lambda X_{p-K+1} + \beta_{m+K}\, Y_m$$

$$+ \gamma_{n+m+K}\, Z_n,$$

$$K = 1, 2, \ldots, p-2 .$$

The general case is treated in a similar manner.

There is an easy way to calculate $X_{L^{**},L}$ just by knowing λ. According to the definition of the eigenvalue λ, we have

$$||X_1,\ldots,X_p,Y_1,\ldots,Y_m,Z_1,\ldots,Z_n||\ L^{**} = \lambda||X_1,\ldots,X_p,Y_1,\ldots,Y_m,Z_1,\ldots,Z_n||$$

which gives the following systems of equations:

$$X_2 = \lambda X_1$$

(12) $$X_3 = \lambda X_2$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$X_p = \lambda X_{p-1}$$

(12.1) $$-\alpha_p X_1 - \alpha_{p-1}\, X_2 - \ldots \ -\alpha_2\, X_{p-1} - \alpha_1\, X_p = \lambda X_p$$

$$Y_2 = \lambda Y_1$$

(13) $$Y_3 = \lambda Y_2$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$Y_m = \lambda Y_{m-1}$$

$$(13.1)\quad (-\beta_{m+p} X_1 - \beta_{m+p-1} X_2 - \dots -\beta_{m+2} X_{p-1} - \beta_{m+1} X_p)$$

$$+ (-\beta_m Y_1 - \beta_{m-1} Y_2 - \dots -\beta_2 Y_{m-1} - \beta_1 Y_m) = \lambda Y_m$$

$$(14)\quad \begin{aligned} Z_2 &= \lambda Z_1 \\ Z_3 &= \lambda Z_2 \\ &\cdots \\ Z_n &= \lambda Z_{n-1} \end{aligned}$$

$$(14.1)\quad (-\gamma_{n+m+p} X_1 - \gamma_{n+m+p-1} X_2 - \dots - \gamma_{n+m+2} X_{p-1} - \gamma_{n+m+1} X_p)$$

$$+ (-\gamma_{n+m} Y_1 - \gamma_{n+m-1} Y_2 - \dots - \gamma_{n+2} Y_{m-1} - \gamma_{n+1} Y_m)$$

$$+ (-\gamma_n Z_1 - \gamma_{n-1} Z_2 - \dots - \gamma_2 Z_{n-1} - \gamma_1 Z_n) = \lambda Z_n.$$

From the system of equations (12), it follows that

$$(15)\quad \begin{aligned} X_2 &= \lambda X_1 \\ X_3 &= \lambda X_2 = \lambda^2 X_1 \\ &\cdots\cdots \\ X_p &= \lambda X_{p-1} = \lambda^{p-1} X_1 . \end{aligned}$$

From (12.1),

$$- \alpha_p X_1 - \alpha_{p-1} X_2 - \dots - \alpha_2 X_{p-1} - \alpha_1 X_p - \lambda X_p = 0 ,$$

$$- \alpha_p X_1 - \alpha_{p-1} \lambda X_1 - \dots - \alpha_2 \lambda^{p-2} X_1 - \alpha_1 \lambda^{p-1} X_1 - \lambda^p X_1 = 0 ,$$

$$- (\alpha_p + \alpha_{p-1}\lambda + \dots + \alpha_2 \lambda^{p-2} + \alpha_1 \lambda^{p-1} + \lambda^p) X_1 = 0 ,$$

and since X_1 is arbitrary it follows that

(15.1) $$\alpha_p + \alpha_{p-1}\lambda + \dots + \alpha_1 \lambda^{p-1} + \lambda^p = 0 .$$

Applying the same technique to (13), we get

(16) $$\begin{aligned} Y_2 &= \lambda Y_1 \\ &\cdots\cdots \\ Y_m &= \lambda Y_{m-1} = \lambda^{m-1} Y_1 \end{aligned}$$

From (13.1),

(16.1) $$(-\beta_{m+p} - \beta_{m+p-1}\lambda - \dots - \beta_{m+1}\lambda^{p-1}) X_1$$

$$+ (-\beta_m - \beta_{m-1}\lambda - \dots - \beta_1 \lambda^{m-1} - \lambda^m) Y_1 = 0.$$

Since X_1, Y_1 are arbitrary real numbers it follows from (16.1) that

(17.1) $$-\beta_{m+p} - \beta_{m+p-1}\lambda - \dots - \beta_{m+1}\lambda^{p-1} = 0$$

and

(17.2) $$-\beta_m - \beta_{m-1}\lambda - \dots - \beta_1\lambda^{m-1} - \lambda^m = 0.$$

Equation (17.2) tells us that $\lambda^m + \beta_1\lambda^{m-1} + \dots + \beta_m = 0$ is another polynomial of λ which can be used to calculate λ.

Equations (15.1) and (17.1) along with the well known result relating the coefficients of the characteristic polynomial with the eigenvalues, give

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = (-1)\ \alpha_1,$$

$$\lambda_1 + \dots + \lambda_{p-1} = (-1)\ \frac{\beta_{m+2}}{\beta_{m+1}} .$$

(18.1) Therefore, $\lambda_p = (-1)\ (\alpha_1 + \frac{\beta_{m+2}}{\beta_{m+1}})$,

Also since, $\lambda_1\lambda_2 \dots \lambda_p = (-1)^p\ \alpha_p$ and

$$\lambda_1 \dots \lambda_{p-1} = (-1)^{p-1}\ \frac{\beta_{m+p}}{\beta_{m+1}}$$

we obtain

$$(18.2) \qquad \lambda_p = -\frac{\alpha_p \beta_{m+1}}{\beta_{m+p}} .$$

We can apply the same technique to (14.1).

To summarize the previous discussion, we can state the following:

Theorem 2.6: If λ is given then the left eigenvector for L** is

$$||X_1, \lambda X_1, \cdots , \lambda^{p-1} X_1, Y_1, \lambda Y_1, \ldots, \lambda^{m-1} Y_1, Z_1, \lambda Z_1, \ldots, \lambda^{n-1} Z_1||$$

where X_1, Y_1, Z_1 are arbitrary nonzero constants.

From (18.1) or (18.2), it follows that

Theorem 2.7: If A has been reduced to L** (but not L) then one of its eigenvalue is $(-1)\left(\alpha_1 - \frac{\beta_{m+2}}{\beta_{m+1}}\right)$ or is $-\frac{\alpha_p \beta_{m+1}}{\beta_{m+p}}$ and its corresponding left eigenvector is the same as the form described in Theorem 2.6.

Now, we can use Method α to calculate the right eigenvector and use Theorem 2.6 to calculate the left eigenvector for L** when λ is given. Theorem 2.6 tells us the easiest way to get $X_{L^{**},L}$, which is the left eigenvector for L**. By Theorem 2.4, the left eigenvector for A is just $X_{L^{**},L}$ (πE), where πE has been calculated while reducing A to L**. The above discussion presents a very easy way to get the left eigenvector for A, but are there some easy ways to get the right eigenvector for A?

The answer is yes since we can use an analogous procedure as in deriving Method α, Theorem 2.6 and Theorem 2.7 to get Method β, Theorem 2.8 and 2.9 as shown below. First we restrict L^* to the following form:

$$\left\|\begin{matrix}
-a_1 & -a_2 & \cdots & -a_s & -a_{s+1} & & \cdots & -a_{s+t} & -a_{s+t+1} & & \cdots & -a_{s+t+j} \\
1 & 0 & & 0 & & & & & & & & \\
0 & 1 & & \vdots & & & & & & & & \\
 & & \ddots \; 1 & 0 & & & & & & & & \\
 & & & & -b_1 & -b_2 & \cdots & -b_t & -b_{t+1} & & \cdots & -b_{t+j} \\
 & & & & 1 & 0 & & 0 & & & & \\
 & & & & 0 & 1 & & \vdots & & & & \\
 & & & & & & \ddots \; 1 & 0 & & & & \\
 & & & & & & & & -c_1 & -c_2 & \cdots & -c_j \\
 & & & & & & & & 1 & 0 & & 0 \\
 & & & & & & & & 0 & 1 & & \vdots \\
 & & & & & & & & & & \ddots \; 1 & 0
\end{matrix}\right\|$$

then we will get results analogous to Theorem 2.5 and Theorem 2.7.

Theorem 2.8: If λ is given, the right eigenvector for L^* is

$$\| \lambda^{s-1} X_s,\ \lambda^{s-2} X_s,\ \ldots, \lambda X_s,\ X_s,\ \lambda^{t-1} Y_t,\ \lambda^{t-2} Y_t,\ \ldots,\ Y_t,\ \lambda^{j-1} Z_j,\ \ldots,\ Z_j \|^T$$

where X_s, Y_t, Z_j are arbitrary nonzero constants.

Theorem 2.9: If A has been reduced to L*, then one of its eigenvalues is $(-1)(b_1 - \frac{a_{s+2}}{a_{s+1}})$ or is $-\frac{b_t\, a_{s+1}}{a_{s+t}}$ and its corresponding right eigenvector is the form described in Theorem 2.8.

Method β, analogous to method α, is to calculate the left-eigenvector of L*.

First, set $X_s = 1$,

$$X_K = \lambda X_{K-1} + a_{K-1}\, X_1 ,$$

$$K = 2, \ldots, s-1.$$

Next, set $Y_t = 1$,

$$Y_1 = -\,(\lambda + a_{s+t}\, X_1)/b_t ,$$

$$Y_K = \lambda Y_{K-1} + b_{K-1}\, Y_1 + a_{s+K-1}\, X_1 ,$$

$$K = 2, \ldots, t-1.$$

Then set $Z_j = 1$,

$$Z_1 = -\,(\lambda + b_{t+j}\, Y_1 + a_{s+t+j}\, X_1)/c_j ,$$

$$Z_K = \lambda Z_{K-1} + c_{K-1}\, Z_1 + b_{t+K-1}\, Y_1$$

$$+ a_{s+t+K-1}\, X_1 ,$$

$$K = 2, \ldots, j-1.$$

Extension to the more general case is obvious.

We omit the proof for Theorem 2.8 and 2.9, since the same approach to prove Theorem 2.6 and 2.7 is used. But we should point out that Theorem 2.8 is the easiest way to get the right eigenvector of L*, and then to get the right eigenvector of A.

Theorem 2.6 and Theorem 2.9 show an easy way to calculate an eigenvalue directly from L* or L**. But unfortunately, we still do not have any advanced methods that will guarantee that A can always be reduced to L* or L** instead of L. This may be a very good topic for further study.

Chapter 3

Numerical Examples

(1) Numerical example for reducing A to L with non-zero pivot elements in the whole reducing process

$$A = \left\|\begin{matrix} 1 & 3 & 2 \\ 4 & 5 & 7 \\ 3 & 8 & 4 \end{matrix}\right\|$$

$$E_2^1 = \left\|\begin{matrix} 1 & -1/4 & 0 \\ 0 & 1/4 & 0 \\ 0 & -3/4 & 1 \end{matrix}\right\| \qquad (E_2^1)^{-1} = \left\|\begin{matrix} 1 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 3 & 1 \end{matrix}\right\|$$

$$E_2^1\ A(E_2^1)^{-1} = \left\|\begin{matrix} 0 & 7.75 & 0.25 \\ 1 & 11.25 & 1.75 \\ 0 & 13.25 & -1.25 \end{matrix}\right\| = A^{(1)}$$

$$E_3^2 = \left\|\begin{matrix} 1 & 0 & -7.75/13.25 \\ 0 & 1 & -11.25/13.25 \\ 0 & 0 & 1/13.25 \end{matrix}\right\| \quad (E_3^2)^{-1} = \left\|\begin{matrix} 1 & 0 & 7.75 \\ 0 & 1 & 11.25 \\ 0 & 0 & 13.25 \end{matrix}\right\|$$

$$E_3^2\ A^{(1)}\ (E_3^2)^{-1} = \left\|\begin{matrix} 0 & 0 & 13 \\ 1 & 0 & 45 \\ 0 & 1 & 10 \end{matrix}\right\| = L_A$$

We can easily write down the characteristic polynomial of A from L_A:

$(-1)^3\ (\lambda^3 - 10\lambda^2 - 45\lambda - 13)$.

(2) Numerical example for reducing A to L, with zero pivot element substituted by other nonzero element.

$$A = \begin{Vmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ -1 & 3 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{Vmatrix} \quad a_{21} = 0 \quad \text{but} \quad a_{31} \neq 0$$

$$SA = \begin{Vmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 3 & 1 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & 0 \end{Vmatrix} \qquad E^1 = E^1_{2,SA}$$

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

$$E^1\ A(E^1)^{-1} = \begin{Vmatrix} 0 & 0 & 3 & 2 \\ 1 & 0 & -3 & -1 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{Vmatrix} = A^{(1)}$$

$$E^2\ A^{(1)}(E^2)^{-1} = \begin{Vmatrix} 0 & 0 & 5 & 2 \\ 1 & 0 & -6 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & -2 & 0 \end{Vmatrix} = A^{(2)}$$

$$E^3\ A^{(2)}\ (E^3)^{-1} = \left\|\begin{matrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 3 \end{matrix}\right\| = A^{(3)}$$

the characteristic polynomial of A is $(-1)^4\ (\lambda^4 - 3\lambda^3 + 6\lambda^2 - 7\lambda + 4)$.

(3) Numerical example for reducing A to L** (i.e. The final matrix with zero element in pivot position and the rest of lower column element are all zero.)

$$A = \left\|\begin{matrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{matrix}\right\|$$

$$E_2^1 = \left\|\begin{matrix} 1 & 1/3 & 0 & 0 \\ 0 & -1/3 & 0 & 0 \\ 0 & 8/3 & 1 & 0 \\ 0 & 15/3 & 0 & 1 \end{matrix}\right\| \qquad (E_2^1)^{-1} = \left\|\begin{matrix} 1 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 8 & 1 & 0 \\ 0 & 15 & 0 & 1 \end{matrix}\right\|$$

$$E_2^1 A (E_2^1)^{-1} = \left\|\begin{matrix} 0 & 1/3 & -2/3 & 1/3 \\ 1 & -8/3 & 5/3 & -4/3 \\ 0 & 16/3 & -31/3 & 20/3 \\ 0 & 8 & -14 & 9 \end{matrix}\right\| = A^{(1)}$$

$$E_3^2 = \left\|\begin{matrix} 1 & 0 & 1/16 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 3/16 & 0 \\ 0 & 0 & -3/2 & 1 \end{matrix}\right\| \qquad (E_3^2)^{-1} = \left\|\begin{matrix} 1 & 0 & -1/3 & 0 \\ 0 & 1 & -8/3 & 0 \\ 0 & 0 & 16/3 & 0 \\ 0 & 0 & 8 & 1 \end{matrix}\right\|$$

$$E_3^2 A^{(1)} (E_3^2)^{-1} = \left\|\begin{matrix} 0 & 0 & -1 & 3/4 \\ 1 & 0 & -3 & 2 \\ 0 & 1 & -3 & 5/4 \\ 0 & 0 & 0 & -1 \end{matrix}\right\| = A^{(2)}$$

$A^{(2)}_{4,3} = 0$ We terminate searching at $(A^{K-1})_{n,n-1}$ and get $L^{**} = A^{(2)}$.

The characteristic polynomial of A is $(-1)^4\ (\lambda^3 + 3\lambda^2 + 3\lambda + 1)(\lambda + 1)$.

(4) Numerical example for reducing A to L*.

$$A = \left\|\begin{matrix} 4 & 1 & 2 \\ 4 & 2 & 6 \\ 2 & 2 & 4 \end{matrix}\right\|$$

$$R_3^2 = \left\|\begin{matrix} 1 & 0 & 0 \\ -1 & 1/2 & -2 \\ 0 & 0 & 1 \end{matrix}\right\| \qquad (R_3^2)^{-1} = \left\|\begin{matrix} 1 & 0 & 0 \\ 2 & 2 & 4 \\ 0 & 0 & 1 \end{matrix}\right\|$$

$$(R_3^2)^{-1}A\ R_3^2 = \begin{Vmatrix} 3 & -1/2 & 0 \\ 10 & 1 & 4 \\ 0 & 1 & 0 \end{Vmatrix} = A^{(1)}$$

$$R_2^1 = \begin{Vmatrix} 0.1 & -0.1 & -0.4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} \qquad (R_2^1)^{-1} = \begin{Vmatrix} 10 & 1 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$$

$$(R_2^1)^{-1}\ AR_2^1 = \begin{Vmatrix} 4 & -4 & -12 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}$$

The characteristic polynomial of A is $(-1)^3\ (\lambda^3 - 4\lambda^2 + 4\lambda + 12)$.

Chapter 4

Additional Discussion

There exists an interesting example which shows us some unproven properties that may be true in general or just a special case.

If $A = \left\|\begin{matrix} 2.6 & -0.2 & 0.2 \\ 0.2 & 2.1 & 0.4 \\ 0.4 & 0.2 & 0.8 \end{matrix}\right\|$, then

$$E^1_{2,A}\ A(E^1_{2,A})^{-1} = \left\|\begin{matrix} 0 & -7.5 & -5 \\ 1 & 5.5 & 2 \\ 0 & 0 & 2 \end{matrix}\right\| .$$

Note: A is reduced to L** (not L).

By interchanging the rows and the columns of matrix A, but still similar to A, we get the elements of the set $K = \{A^1 \mid A^1 = SAS\}$ where S is an elementary permutation matrix.

Applying our reducing method to set K, the results of $(\pi E)A^1(\pi E)^{-1}$, where $A^1 \varepsilon K$, will eventually result in one of these three forms of L**:

$$\left\|\begin{matrix} 0 & -5 & 2.5 \\ 1 & 4.5 & -1 \\ 0 & 0 & 3 \end{matrix}\right\| ,$$

$$\begin{Vmatrix} 0 & -6 & -1 \\ 1 & 5 & 0.5 \\ 0 & 0 & 2.5 \end{Vmatrix},$$

or

$$\begin{Vmatrix} 0 & -7.5 & -1.5 \\ 1 & 5.5 & 0.5 \\ 0 & 0 & 2 \end{Vmatrix}.$$

This tells us that we can find all the eigenvalues of A simultaneously. But the result of $(\pi R)^{-1} A^1 (\pi R)$ gives only one case $A^1 \varepsilon K$:

$$\begin{Vmatrix} 7.5 & -18.5 & 15 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}.$$

The above results may suggest us that if A can be reduced into L** (or L*) but not L then A^1 also can be reduced to one of the forms of L**(or L*) but not L.

More generalized results are

i) $(\pi_A E)\, A(\pi_A E)^{-1} = L_A^{**}$ but not L_2

if and only if $(\pi_{A^1} E)\, A^1\, (\pi_{A^1} E)^{-1} = L_{A^1}^{**}$

ii) $(\pi_A R)^{-1}\, A(\pi_A R) = L_A^{*}$ but not L_1

if and only if $(\pi_{A^1} R)^{-1}\, A^1\, (\pi_{A^1} R) = L_{A^1}^{**}$

where $A^1 \varepsilon S_n$,

S_n is the set of all permutation matrices of rank n, and

A (or A^1) under π indicates πE, πR is derived from A (or A^1).

If the above statements are not always true, then we can conclude that

$$L^{**} = (\underset{A}{\pi}E)\ (\underset{A^1}{\pi E})^{-1}\ L(\underset{A^1}{\pi E})(\underset{A}{\pi}E)^{-1}.$$

This means that L is similar to L** which seems like it contradicts the fact that a "derogatory" matrix will not be similar to a non-derogatory matrix.

Furthermore, it tells us that the roots (i.e. eigenvalues) of a polynomial (characteristic polynomial) can be solved by a set of similarity transformations, since λ is almost explicitly shown in some coefficients in L** (or L*). But this seems like it is impossible, unless we can prove i) or ii) is wrong.

If it worthwhile to mention here that another two kinds of companion matrix, L_3 or L_4, and their associate "analogous" companion matrices L*** and L**** can be derived from the same discussion.

Let us show it briefly

$$L_3 = \left\| \begin{matrix} P_1 & 1 & 0 & \dots & 0 \\ P_2 & 0 & 1 & \dots & 0 \\ . & . & . & \dots & . \\ P_{n-1} & 0 & 0 & \dots & 1 \\ P_n & 0 & 0 & \dots & 0 \end{matrix} \right\| \qquad L_4 = \left\| \begin{matrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & 1 \\ P_n & P_{n-1} & P_{n-2} & \dots & P_1 \end{matrix} \right\|$$

$$L^{***} = \left\| \begin{array}{cccccc} L_3^{r_1} & & & & & 0 \\ C_{12}^T & L_3^{r_2} & & & & \\ C_{13}^T & C_{23}^T & D_{r_3}^T & & & \\ C_{14}^T & C_{24}^T & F_{34}^T & L_3^{r_4} & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ C_{1s}^T & C_{2s}^T & F_{3s}^T & C_{4s}^T & \cdot & L_3^{r_s} \end{array} \right\|$$

and

$$L^{****} = \left\| \begin{array}{cccccc} L_4^{r_1} & & & & & 0 \\ H_{12}^T & L_4^{r_2} & & & & \\ F_{13}^T & F_{23}^T & D_{r_3}^T & & & \\ H_{14}^T & H_{24}^T & H_{34}^T & L_4^{r_4} & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ H_{1s}^T & H_{2s}^T & H_{3s}^T & H_{4s}^T & \cdot & L_4^{r_s} \end{array} \right\| .$$

Now, if we consider the sequence of pivots in the order (1,2), (2,3), ..., (n-1,n) for defining E_K^{K+1} in Theorem 2.3, then A will be reduced to L***. If we consider the sequence of pivots in the order (n-1,n) (n-2,n-1), ..., (1,2) for defining R_K^{K+1} in Theorem 2.3', then A will be reduced to L****.

Let us look back at Chapter 2 to determine why it is necessary to require a non-trivial reducing process when we determine the eigenvector of L** (or L*). The reason is that (17.1) may not be true unless L** is determined from a nontrivial reducing process, that is, $\pi E \neq I$ or S and $\pi E\ A(\pi E)^{-1} \neq A$. This follows since in the matrix A is originally in the form L** then the elements $\beta_{m+1}, \beta_{m+2}, \ldots, \beta_{m+\rho}$ are arbitrary and need not satisfy (17.1). However, a nontrivial reducing process will guarantee that (17.1) is satisfied. A similar argument holds if A is originally in the form L*. Hence, (17.1), (18.1), (18.2), Theorem 2.7, and Theorem 2.9 may not be true if L** (or L*) is derived by a trivial reducing process. For example. Consider the example in this chapter and another matrix B given by

$$B = \left\| \begin{vmatrix} 0 & -7.5 & 18 \\ 1 & 5.5 & 3 \\ 0 & 0 & 1 \end{vmatrix} \right\| .$$

Then $\pi E = I$

and $\pi EB(\pi E)^{-1} = B$

and Theorem 2.7

tell us that

$$-\left(\alpha_1 - \frac{\beta_3}{\beta_2}\right) = \lambda \qquad \text{But } -\left(-5.5 - \frac{18}{3}\right)$$

$$= 11.5 \neq 2,\ 3 \text{ or } 2.5.$$

thus contradicting Theorem 2.7.

But Three forms for L** for A given in this chapter show that

$$- (\alpha_1 - \frac{\beta_3}{\beta_2}) = - (- 5.5 - \frac{-5}{2}) = 2.5$$

$$= - (- 4.5 - \frac{2.5}{-1}) = 2$$

$$= - (- 5 \quad - \frac{-1}{0.5}) = 3$$

and hence Theorem 2.7 holds.

To summarize, the fundamental theorem for general matrices may be expressed as the following:

Every matrix A may be reduced by a similarity transformation to the matrix L*, L**, L*** and L****.

REFERENCE

[1] Wilkinson, J. H. The Algebraic Eigenvalue Problem, Clarendon Press. Oxford, 1969.

[2] Hammarling, S. J. Latent Roots and Latent Vector, The University of Toronto Press, 1970.

APPENDIX

```
/*  THIS PROGRAM IS TO CALCULATE THE CHARACTERISTIC */
                     /*  POLYNOMIAL OF ANY SQUARE MATRIX */
MAXTRS:     PROC OPTIONS (MAIN);
            DCL (A,B) (20,20) FLOAT BIN INIT((400) 0),
                (X,Y,Z) BIN FIXED,
                (C,TEMP) (20) FLOAT BIN INIT((20) 0),
                (TEMPROW,TEMPCOL) (20) FLOAT BIN,
                TITLE CHAR(29) EXTERNAL;
READIN:     GET LIST (N,((A(I,J) DO J=1 TO N)
                      DO I=1 TO N));
            IF N=0 THEN GOTO ENDFIL;
            PUT PAGE    EDIT ('MATRIX BEFORE TRANSFORMATION')
                             (X(10),A);
                        /* ECHO CHECK FOR INPUT */
            DO I=1 TO N;
            PUT SKIP(2) EDIT ((A(I,J) DO J=1 TO N))
                             (F(10,6),X(2));
            END;
            TITLE='TIME BEFORE TRANSFORMATION : ';CALL TIMER;
LOOP1:      DO I=1 TO N-1;
            J=I+1;
            /* SET AN IDENTITY MATRIX */
            DO X=1 TO N;
            DO Y=1 TO N ; B(X,Y)=0 ; END ;
            B(X,X)=1;
            END ;
            /* CHECK AND AVOID ZERO PIVOT ELEMENT */
            IF A(J,I) > -1.E-05
            & A(J,I) < 1.E-05  THEN GOTO ZERO ;
MULTIP:     DO K=1 TO N;
            /* SET MATRIX B TO BE MODIFIED GAUSS ELIMINATION OPERATOR E(I/J) */
            IF K=J THEN
            B(K,J)=1/A(J,I)  ;
            ELSE B(K,J)=-A(K,I)/A(J,I);
            C(K)=A(K,I);    /* C(K) IS LOADED WITH K-TH COLUMN OF A */
            END MULTIP;
LOOP2:      DO Z=1 TO N;
LOOP3:      DO Y=1 TO N;
LOOP4:      DO X=1 TO N;    /* CALCULATE E(J/K) A */
            TEMP(Y)=TEMP(Y)+B(Y,X)*A(X,Z);
            END LOOP4;
            END LOOP3;
            DO X=1 TO N;
            A(X,Z)=TEMP(X);
            END ;   TEMP=0;
            END LOOP2;
            DO K=1 TO N;
            B(K,J)=C(K);
            END;
LOOP22:     DO Z=1 TO N;
LOOP33:     DO Y=1 TO N ;
LOOP44:     DO X=1 TO N;
```

SOURCE STATEMENT

```
                    /* CALCULATE (E_K^j A(E_K^j)^-1) */
          TEMP(Y)=TEMP(Y)+A(Z,X)*B(X,Y);
          END LOOP44; END LOOP33;
          DO X=1 TO N;
          A(Z,X)=TEMP(X);
          END ;   TEMP=0 ;
                END LOOP22;
          GOTO ECHPUT;
ZERO:     DO;
          J=J+1;
          IF J>N THEN GOTO NEXCOL ;
          /* CHECK THE REST ELEMENT IN THE I-TH COLUMN IS NONZERO */
          IF A(J,I) > 1.E-05
          & A(J,I) < 1.E-05 THEN GOTO NEXROW;
          /* IF NONZERO ELEMENT OCCURS, A BE INTERCHANGED ITS COLUMNS
            AND ROWS */
          DO L=1 TO N;
          TEMPROW(L)=A(J,L);
          A(J,L)=A(I+1,L);
          A(I+1,L)=TEMPROW(L);
          END;
          DO L=1  TO N;
          TEMPCOL(L)=A(L,J);
          A(L,J)=A(L,I+1);
          A(L,I+1)=TEMPOOL(L);
          END;
          J=I+1 ;
          GOTO MULTIP;
NEXROW:   END ZERO;
ECHPUT:   PUT SKIP(2) ;
          IF I=N-1 THEN GOTO NEXCOL :
          /* PRINT THE INTERMEDIATE RESULT */
          DO X=1 TO N ;
          PUT SKIP(2) EDIT ((A(X,Y) DO Y=1 TO N))
                           (F(15,6),X(2));
          END;
NEXCOL:   END LOOP1;
          TITLE='TIME AFTER TRANSFORMATION : ';
          CALL TIMER;
          /* PRINT THE FINAL RESULTS */
          PUT SKIP(3) EDIT ('MATRIX AFTER TRANSFORMATION')
                        (X(10),A);
          DO I=1 TO N;
          PUT SKIP(2) EDIT ((A(I,J) DO J=1 TO N))
                         (F(15,6),X(2));
          END;
          /* PROCEDURE FOR INFORMATIONS ABOUT TIME */
TIMER:    PROC ;
          DCL T CHAR(9),SECT(4) CHAR(3),TSTRING CHAR(12),
              TIME BUILTIN, TITLE CHAR(29) EXTERNAL,
              (I,K) FIXED;
```

```
T=TIME;
K=-1;
DO I=1 TO 3 ;
K=K+2;
SECT(I)=SUBSTR(T,K,2)  || '.' ;
END;
SECT(4)=SUBSTR(T,7,3);
TSTRING=SECT(1) || SECT(2) || SECT(3) || SECT(4);
PUT SKIP(2) EDIT (TITLE,TSTRING)
                 (X(30),A,X(3),A);
PUT SKIP; END; GOTO READIN;
```

GAUSS ELIMINATION OPERATOR FOR

DETERMINING THE CHARACTERISTIC POLYNOMIAL

AND EIGENVECTORS OF ANY SQUARE MATRIX

by

CHI-CHING CHEN

B.S., NATIONAL TSING HUA UNIVERSITY, 1969

AN ABSTRACT OF A MASTER'S THESIS

Submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1972

ABSTRACT

An approach which uses a modified Gauss elimination operator as a successive similarity transformation, transforms any square matrix into its companion matrix in n-1 steps and gets its characteristic polynomial.

An algorithm for finding for eigenvectors is developed when the eigenvalues are given.

Numerical examples and a computer program written in PLC are provided to illustrate the method.