A NOTE ON THE SOLUTION OF THE UNILATERAL MATRIX EQUATION

JAMES H. BELL

M. H. Ingraham [1][1] has developed an algorithm for the solution of the unilateral matrix equation $\sum_{i=0}^{r} R_i X^i = 0$, where the coefficients R_i $(i=0, 1, 2, \cdots, n)$ are $n \times n$ matrices with elements in a field $\mathcal{J}$ of characteristic zero. It is the purpose of this note to extend his algorithm to include the unilateral equation in which the coefficients are $m \times n$ matrices. It is shown that the solution in this case is easily reduced to the solution of an equation with $n \times n$ matrix coefficients. If in particular $m < n$, it is shown that if the equation has one solution it has an infinitude of solutions.

Ingraham's algorithm is based upon the following facts:

(1). The matrix X is a solution of $\sum_{i=0}^{r} R_i X^i = 0$ if and only if the *canonical triangular form* (c.t.f.)[2] A of $\Lambda - X$ $(\Lambda = \lambda I^{n \times n}$ where λ is a commutative indeterminate) is a right divisor of the c.t.f. $P(\Lambda)$ of $\sum_{i=0}^{r} R_i \Lambda^i$. (In the remainder of this paper we shall use $R(X)$ to represent $\sum_{i=0}^{r} R_i X^i$, $R(\Lambda)$ to represent $\sum_{i=0}^{r} R_i \Lambda^i$, and so forth.)

(2). The problem of factoring $P(\Lambda) = Q(\Lambda) A(\Lambda)$ is reduced to solving equations of the type $p_{ii} = q_{ii} a_{ii}$, $q_{ii} a_{ij} \equiv p_{ij} - \sum_{l=i+1}^{j-1} q_{il} a_{lj}$ mod a_{jj}, and $q_{ji} = a_{ji} = 0$ $(j > i,\ i = 1, 2, \cdots, n)$, where the elements a_{ii} are so chosen that the degree of $\prod_{i=1}^{j} a_{ii}$ is less than or equal to j, and the degree of $\prod_{i=1}^{n} a_{ii}$ is equal to n.

(3). The necessary and sufficient condition that $A(\Lambda) = \sum_{i=0}^{s} A_i \Lambda^i$ be the c.t.f. of a matrix $\Lambda - X$ is that $W_1 = (A_s, A_{s-1}, \cdots, A_1)$ be of rank n [3].

Since X is of necessity square of order n, fact 3 is still valid. The first fact is based upon a factor theorem, namely, X is a solution of $R(X) = 0$ if and only if $R(\Lambda) = S(\Lambda)(\Lambda - X)$. The dimensions of the coefficients do not affect this factor theorem. However, steps 1 and 2 involve the concept of the c.t.f. of $R(\Lambda)$ and must be reconsidered. The cases $m > n$ and $m < n$ will be treated separately.

Case I. $m > n$. Let

Presented to the Society, February 24, 1950; received by the editors June 20, 1950.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] The canonical triangular form of a matrix with elements in $\mathcal{J}[\lambda]$ is a matrix with zeros below the main diagonal. The main diagonal elements are monic polynomials or zero. If a main diagonal element is zero, the row in which it occurs consists entirely of zeros. The elements of any column are reduced modulo the main diagonal element of that column.

$$R(\Lambda) = \begin{pmatrix} P_1^{n\times n} \\ P_2 \end{pmatrix}^{m\times n}.$$

Proceeding in the same manner as is used to find the c.t.f. of a square matrix,[3] it may be shown that there is a unimodular matrix $U^{m\times m}$ such that

$$UR(\Lambda) = \begin{pmatrix} P^{n\times n} \\ 0 \end{pmatrix}^{m\times n},$$

where P is unique and is in canonical triangular form. If X is a solution of $R(X) = 0$, it follows that

$$UR(\Lambda) = \begin{pmatrix} Q_1^{n\times n} \\ Q_2 \end{pmatrix}(\Lambda - X) = \begin{pmatrix} P \\ 0 \end{pmatrix} = \begin{pmatrix} Q^{n\times n} \\ 0 \end{pmatrix} A,$$

where A is the c.t.f. of $\Lambda - X$. Likewise, if $P = QA$, it follows that

$$R(\Lambda) = U^{-1}\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}(\Lambda - X),$$

and the following theorem is true.

THEOREM 1. *If $m > n$, X is a solution of $R(X) = 0$ if and only if X is a solution of $P(X) = 0$ where $P(X)$ has $n\times n$ matrices as coefficients.*

Let $R_n(\Lambda)$ be an $n\times n$ matrix formed by using n rows of $R(\Lambda)$.

LEMMA 1. *If $R(X) = 0$ has a solution and if $|R_n(\Lambda)|$ is considered for all possible choices of the n rows, then these determinants must have a common factor of degree n.*

PROOF. Since $R(\Lambda) = S(\Lambda)(\Lambda - X)$, it follows that $R_n(\Lambda) = S_n(\Lambda)(\Lambda - X)$, and

$$|R_n(\Lambda)| = |S_n(\Lambda)| \cdot |\Lambda - X|.$$

This result is given by Roth [4] in the case that $\mathcal{J}$ is the complex field. The proof above is somewhat simpler.

Case II. $m < n$. Let $R(\Lambda) = (P_1^{m\times m}, P_2)^{m\times n}$ and form the matrix

$$R^* = \begin{pmatrix} P_1 & P_2 \\ 0 & 0 \end{pmatrix}^{n\times n}.$$

THEOREM 2. *The matrix X is a solution of $R(X) = 0$ if and only if X*

[3] The canonical triangular form is obtained in the same manner as the Hermite normal form [2] except that the columns are operated on in reverse order.

is a solution of $R^*(X)=0$. *If there exists one solution of* $R(X)=0$, *there exists an infinitude of solutions.*

There exists a unimodular matrix T^* such that

$$T^*R^* = \begin{pmatrix} P_{11} & P_{12} \\ 0 & P_{22} \end{pmatrix} = P^*$$

where P^* is the c.t.f. of R^*, P_{11} is the c.t.f. of P_1, and $P_{22}=0$ if P_1 is nonsingular.

If X is a solution of $R(X)=0$, then $R(\Lambda)=(P_1, P_2)=(S_1,S_2)(\Lambda-X)$, and

$$R^*(\Lambda) = \begin{pmatrix} S_1 & S_2 \\ 0 & 0 \end{pmatrix}(\Lambda - X).$$

Also, if

$$R^*(\Lambda) = \begin{pmatrix} P_1 & P_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}(\Lambda - X),$$

it is easily shown that $B_3=0$, $B_4=0$, and therefore $R(\Lambda)=(P_1, P_2)$ $=(B_1, B_2)(\Lambda-X)$.

Therefore, X is a solution of $R(X)=0$ if and only if X is a solution of the matrix equation $R^*(X)=0$, whose coefficients are $n\times n$ matrices. However, since $|R^*(\Lambda)|=0$ it follows that if $R^*(X)=0$ has a solution, $R^*(X)=0$ and consequently $R(X)=0$ have an infinite number of solutions [5].

THEOREM 3. *If* P_l^* *is the* $l\times l$ *matrix formed by the first* l *rows and* l *columns of* P^*, *where* l *is chosen such that* p_{ll} *is the last nonzero main diagonal element of* P^*, *then if* $P_l^*(X)=0$ *has a solution*, $R(X)=0$ *has a solution.*

Suppose $P_l^*=\overline{Q}_l(\Lambda_l-X_l)=Q_lA_l^{l\times l}$ where $T_lA_l=(\Lambda_l-X_l)$ (T_l unimodular). Let

$$A^{n\times n} = \begin{bmatrix} A_l & \cdot & \cdots & \cdot \\ 0 & \lambda - c_1 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \lambda - c_{n-l} \end{bmatrix},$$

where the c_j are arbitrarily selected so that $\lambda-c_j$ and p_{ii} are relatively prime for all i, $j=1, 2, \cdots, n-l$, and $a_{is}=0$ $(i>s)$. Then $A^{n\times n}$ can be shown to be a right divisor of $P^*(\Lambda)$ where the elements a_{is} $(i<s,\ s=l+1, l+2, \cdots, n)$ will be constants and will be uniquely

determined in terms of the elements c_j. The matrix W_1 connected with A, since the rank of $(W_1)_l$ is l, will be of rank $l+n-l=n$. That is, A is the left associate of a matrix of the form $\Lambda - X$, and $X^{n\times n}$ will be a solution of $R(X)=0$.

As a special case of the above result, if $|P_1|=0$ and $R(X)=0$ has a solution, there will exist an infinite family of solutions.[4]

The above results have an immediate application to the solution of the equation $\sum_{m=0}^{s} A_m \times (K_m X^m)=0$ where $K_m=(k_{m,ij})^{t\times n}$, $A_m=(a_{m,ij})^{r\times p}$, $A\cdot\times B=A\times B=(Ab_{ij})$ [2], and 0 is the $rt\times pn$ zero matrix.

In a paper submitted to the Proceedings [6], it is shown that the equation $Q(X)=\sum_{m=0}^{s} A_m\times(K_m X^m)=0$ has a solution if and only if the unilateral matrix equations $\sum_{m=0}^{s} a_{m,ij}K_m X^m=0$ $(i=1, 2, \cdots, r;$ $j=1, 2, \cdots, p)$ have a common solution. It is also shown that these equations, pr in number, may be reduced to an equivalent set of equations, q in number, equal to the number of linearly independent (over $\mathfrak{J}$) matrices in the set $A_0, A_1, \cdots, A_s$. That is, if X is a solution of $Q(X)=0$, X is a solution of $q\leqq s+1$ equations of the form $\sum_{m=0}^{s} a_{m,k}K_m X^m=0$ $(k=1, 2, \cdots, q)$. Therefore X is a solution of

$$\sum_{m=0}^{s} R_m X^m = \sum_{m=0}^{s} \begin{pmatrix} a_{m,1} & K_m \\ a_{m,2} & K_m \\ \cdot & \\ \cdot & \\ \cdot & \\ a_{m,q} & K_m \end{pmatrix} X^m = 0, \tag{1}$$

where R_m is a $tq\times n$ matrix, and 0 is the $tq\times n$ zero matrix. Likewise if X is a solution of (1), X is a solution of $\sum_{m=0}^{s} A_m\times(K_m X^m)=0$.

THEOREM 4. *The unilateral direct product matrix equation* $\sum_{m=0}^{s} A_m \times(K_m X^m)=0$ *has a solution if and only if the equation* (1) *has a solution.*

The solution of the equation (1) falls into case I or case II according as $tq>n$ or $tq<n$.

COROLLARY. *If* $tq<n$ *and there is one solution of* $\sum_{m=0}^{s} A_m\times(K_m X^m)=0$, *there is an infinitude of solutions.*

BIBLIOGRAPHY

1. M. H. Ingraham, *Rational methods in matrix equations*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 61–70.

[4] A family of solutions is a set of solutions X_j such that the canonical triangular forms of the $\Lambda - X_j$ have the same main diagonal [5].

2. C. C. MacDuffee, *The theory of matrices*, Ergebnisse der Mathematik und ihrer Grenzgebiete, New York, Chelsea, 1946, pp. 31–36.

3. J. H. Bell, *Left associates of monic matrices, with an application to unilateral matrix equations*, Amer. J. Math. vol. 71 (1949).

4. W. E. Roth, *On the unilateral equation in matrices*, Trans. Amer. Math. Soc. vol. 32 (1930) p. 67.

5. J. H. Bell, *Families of solutions of the unilateral matrix equation*, Proceedings of the American Mathematical Society vol. 1 (1950) pp. 152, 157.

6. ———, *The solution of a unilateral direct product matrix equation*, Proceedings of the American Mathematical Society vol. 1 (1950) pp. 777–781.

MICHIGAN STATE COLLEGE

ON MAGIC SQUARES CONSTRUCTED BY THE UNIFORM STEP METHOD

T. M. APOSTOL AND HERBERT S. ZUCKERMAN

An application of the theory of congruences to the study of magic squares constructed by the uniform step method was first given by D. N. Lehmer.[1] The n^2 cells of the square are denoted by two coordinates (A, B), A being the number of the column counting from the left and B the number of the row counting from the bottom. Lehmer summed up the uniform step process in the following congruences for determining the cell (A_x, B_x) into which the number x is entered:

$$A_x \equiv p + \alpha(x-1) + a\left[\frac{x-1}{n}\right] \pmod{n}, \tag{1}$$

$$B_x \equiv q + \beta(x-1) + b\left[\frac{x-1}{n}\right] \pmod{n}, \tag{2}$$

where (p, q) is the cell into which the number 1 is entered, (α, β) is the "step" used in proceeding from one cell to another, (a, b) is the "break-step" that must be used when an occupied cell is arrived at, and the symbol $[k]$ denotes the greatest integer contained in k. Lehmer proved the following theorems:

Presented to the Society, November 25, 1950; received by the editors August 21, 1950.

[1] D. N. Lehmer, *On the congruences connected with certain magic squares*, Trans. Amer. Math. Soc. vol. 31 (1929) pp. 529–551. Definitions of the terms "magic," "diabolic," and "symmetric" are given in this paper.