THE SOLUTION OF A UNILATERAL DIRECT PRODUCT MATRIX EQUATION

JAMES H. BELL

Introduction. In the study of families of solutions of the unilateral matrix equation, over a field $\mathcal{J}$ of characteristic zero, the matrix equation $\sum_{m=0}^{s} A_m \cdot \times X^m = 0$ occurred. Necessary and sufficient conditions were obtained for the solution of this equation.[1] The more general equation $\sum_{m=0}^{s} A_m \cdot \times (K_m X^m) = 0$ is considered in this paper. The solution of this general equation is reduced to the solution of a system of simultaneous unilateral matrix equations, which are equal in number to the number of linearly independent (over $\mathcal{J}$) matrices in the set $A_s, A_{s-1}, \cdots, A_0$. Also, it is shown how, under certain conditions, the equation may possibly be solved using a method paralleling the method of M. H. Ingraham [1][2] for the solution of the unilateral matrix equation. In this connection, an interesting remainder theorem and a divisor theorem are obtained.

In this paper the matrices involved in the equations will have elements in $\mathcal{J}$. For the sake of brevity $A \cdot \times B = (Ab_{ij})$ will be written $A \times B$. Let λ be a scalar indeterminate. A matrix with elements in the polynomial ring $\mathcal{J}[\lambda]$ will be termed a *λ-matrix*. A square *λ-matrix* which has an inverse which is also a λ-matrix is called *unimodular*. If $TA = B$ where T, A, and B are λ-matrices and T is unimodular, then A is said to be a *left associate* of B.

Every square λ-matrix is the left associate of a unique λ-matrix of the following form: Every element below the main diagonal is zero. If the main diagonal element is nonzero, it is a monic polynomial and the other elements in its column are of lesser degree in λ. In the case that the main diagonal element is zero, then all of the elements in its row are also zero. This matrix will be referred to as the *canonical triangular form or c.t.f.*[3]

1. **The general equation.** In this section X will be considered to be an $n \times n$ matrix, A_m $(m = 0, 1, \cdots, s)$ will be an $r \times p$ matrix,

---

Presented to the Society, September 1, 1949; received by the editors September 23, 1949.

[1] J. H. Bell, *Families of solutions of the unilateral matrix equation*, Proceedings of the American Mathematical Society vol. 1 (1950) pp. 151–159.

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

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

and K_m will be $t\times n$. If $M=A^{r\times p}\times B^{t\times n}=(m_{ij})^{rt\times pn}$, the matrix $[M]_{ij}$ will be defined to be the matrix $(m_{(k-1)r+i,(l-1)p+j})$, $k=1,2,\cdots,t$; $l=1,2,\cdots,n$. It follows that $[M]_{ij}=a_{ij}B$, $[M]_{ij}+[N]_{ij}=[M+N]_{ij}$, and $[M]_{ij}\equiv 0$ for all i and j, if and only if $M\equiv 0$.

THEOREM 1. *The equation* $\sum_{m=0}^{s} A_m\times(K_mX^m)=0$ *over* $\mathfrak{J}$, *where* $A_m=(a_{m,ij})^{r\times p}$, *has a solution if and only if the unilateral matrix equations* $\sum_{m=0}^{s} a_{m,ij}K_mX^m=0$ $(i=1,2,\cdots,r;\ j=1,2,\cdots,p)$ *have a common solution. These simultaneous equations may be reduced in number to an equivalent set equal in number to the number of linearly independent (over* $\mathfrak{J}$*) matrices in the set* $A_s, A_{s-1},\cdots,A_0$.

Let $Q=\sum_{m=0}^{s} A_m\times(K_mX^m)$, then $[Q]_{ij}=\sum_{m=0}^{s} a_{m,ij}K_mX^m$. Therefore, $Q=0$ if and only if $[Q]_{ij}=0$ for every i and j. That is, X is a solution of $Q=0$ if and only if X is a solution of the simultaneous set of equations $\sum_{m=0}^{s} a_{m,ij}K_mX^m=0$.

Since each equation is of degree s, and there are rp equations, it may be possible to pick out a basic set of linearly independent equations. Any solution of the basic equations will be a solution of the whole set. The scalar coefficients (the $a_{m,ij}$) of any equation will be linear combinations of the coefficients of the basic equations. Form the matrix $(a_{s,ij}, a_{s-1,ij},\cdots,a_{0,ij})$ $(i=1,2,\cdots,r; j=1,2,\cdots,p)$. The rank of this matrix will equal the number of linearly independent equations and will also equal the number of linearly independent columns in the matrix. Each column is one of the matrices A_m considered as a vector in rp space. Hence the number of linearly independent equations equals the number of linearly independent matrices among $A_s, A_{s-1},\cdots,A_0$.

THEOREM 2. *If the matrix* K_m $(m=0,1,\cdots,s)$ *is square and of order* n, *the equation* $\sum_{m=0}^{s} A_m\times(K_mX^m)=0$ *will have a solution only if the* rp *polynomials formed by taking the determinant of* $\sum_{m=0}^{s}a_{m,ij}K_m\lambda^m$, *for every* i *and* j, *have a common divisor of degree* n.

If X is a solution of $\sum_{m=0}^{s} a_{m,ij}K_mX^m=0$, then the characteristic polynomial of X must divide the determinant of $\sum_{m=0}^{s} a_{m,ij}K_m\lambda^m$ [1].

COROLLARY. *The equation* $\sum_{m=0}^{s} A_m\times X^m=0$ *has a solution if and only if the polynomials* $\sum_{m=0}^{s} a_{m,ij}\lambda^m$ *have a common divisor* $\sum_{m=0}^{q} d_m\lambda^m$ *which is not a unit, and is such that* $\sum_{m=0}^{q} d_mX^m=0$ *has a solution.*

2. **An alternative method.** If all of the matrices are square although the order of A_m may be different from the order of X, the solution of the equation may be carried out by a method paralleling Ingraham's method for the solution of the unilateral equation [1].

Let $\Lambda=\lambda I^{n\times n}$ and form the matrix $\sum_{m=0}^{s} A_m\times(K_m\Lambda^m)$. This matrix will be a λ-matrix. It is a well known theorem that $(A\times B)(C\times D)=(AC\times BD)$, if A and C, and B and D, are of the same orders.

THEOREM 3. *If A and B are two square λ-matrices, not necessarily of the same order, and if A^H and B^H are the canonical triangular forms of A and B respectively, then $(A\times B)^H=A^H\times B^H$. If $T_1A=A^H$ and $T_2B=B^H$, then $(T_1\times T_2)(A\times B)=(A\times B)^H$.*

Suppose A and B are λ-matrices and the c.t.f.'s of A and B are A^H and B^H respectively where $T_1A=A^H$ and $T_2B=B^H$. It follows that $(T_1\times T_2)(A\times B)=A^H\times B^H$. An inspection of $A^H\times B^H$ will show that it is in canonical triangular form. Therefore, $A^H\times B^H=(A\times B)^H$, since $(T_1\times T_2)$ is unimodular.

THEOREM 4. *If $Q=\sum_{m=0}^{s} A_m\times(K_m\Lambda^m)$, then $Q=S(I\times(\Lambda-X))+\sum_{m=0}^{s}A_m\times(K_mX^m)$, where I is of the same order r as A_m, and S is an $rn\times rn$ λ-matrix.*

The proof of Theorem 4 is by induction on s since

$$A_1\times(K_1\Lambda)+A_0\times K_0 = (A_1\times K_1)(I\times(\Lambda-X))+A_1\times(K_1X)+A_0\times K_0$$

and

$$\sum_{m=0}^{s} A_m\times(K_m\Lambda^m) = (A_s\times(K_s\Lambda^{s-1}))(I\times(\Lambda-X))+A_s\times(K_sX\Lambda^{s-1}) + \sum_{m=0}^{s-1} A_m\times(K_m\Lambda^m).$$

COROLLARY. *The matrix X is a solution of $\sum_{m=0}^{s}A_m\times(K_mX^m)=0$ if and only if $\sum_{m=0}^{s} A_m\times(K_m\Lambda^m)=S(I\times(\Lambda-X))$.*

This corollary follows directly from Theorem 4.

Since the matrix X is a solution of $\sum_{m=0}^{s} A_m\times(K_mX^m)=0$ if and only if $Q=\sum_{m=0}^{s} A_m\times(K_m\Lambda^m)=S(I\times(\Lambda-X))$, then X is a solution if and only if $Q^H=R(I\times(\Lambda-X))$. Let $A=(\Lambda-X)^H$, then $TA=\Lambda-X$ where T is unimodular. Also, $(I\times(\Lambda-X))^H=I\times A$. Therefore, X is a solution if and only if there exists a matrix A in canonical triangular form such that A is the left associate of $\Lambda-X$, that is, $TA=\Lambda-X$, and $Q^H=R(I\times T^{-1})(I\times T)(I\times(\Lambda-X))=P(I\times A)$.

The problem of solving the original equation is thus reduced to that of factoring Q^H so as to obtain a right divisor which is of the form $(I\times A)$, where A is the c.t.f. of a matrix $\Lambda - X$ and X has elements in $\mathcal{J}$.

A matrix A is the c.t.f. of a matrix $\Lambda - X$ if: The degree of $\prod_{m=1}^{i} a_{mm}$ is less than or equal to i, the degree of $\prod_{m=1}^{n} a_{mm}$ is equal to n, and if $A=\sum_{m=0}^{p} B_m\Lambda^m$ the matrix $W_1 = \|B_p, B_{p-1}, \cdots, B_1\|$ is of rank n. These conditions are necessary and sufficient [3].

Let $Q^H = (q_{ij})$. Since Q^H and $I\times A$ are triangular matrices, it follows that P is a triangular matrix and

$$q_{(i-1)r+t,(i-1)r+t} = p_{(i-1)r+t,(i-1)r+t}a_{ii} \qquad (t = 1, 2, \cdots, r). \tag{1}$$

Thus a_{ii} $(i=1, 2, \cdots, n)$ must be picked to satisfy (1) and the conditions listed above.

The other elements of the matrix A are obtained in the following manner: In any column, say the jth: a_{jj}, $a_{j-1,j}$, $\cdots$, $a_{1,j}$ are obtained in the order listed, and the columns are obtained in the order $j=1, 2, \cdots, n$. The matrix $I\times A$ is composed of diagonal matrices A_{ij} of order r, and $A_{ij} = a_{ij}I^{r\times r}$.

Assume that all of the elements up to a_{li} have been found, then all p_{kj} $(j\leqq(i-1)r;\ k>lr,\ j=(i-1)r+1,\ (i-1)r+2, \cdots, ir)$ are determined. The following system of equations determines a_{li}:

$$p_{(l-1)r+t,(l-1)r+j} = 0, \qquad t>j;$$

and

$$\sum_{m=l}^{i} p_{(l-1)r+t,(m-1)r+j}a_{mi} = q_{(l-1)r+t,(i-1)r+j}, \qquad 1\leqq t\leqq j\leqq r.$$

That is,

$$p_{(l-1)r+t,(l-1)r+j}a_{li} \equiv q_{(l-1)r+t,(i-1)r+j} - \sum_{m=l+1}^{i-1} p_{(l-1)r+t,(m-1)r+j}a_{mi} \bmod a_{ii}. \tag{2}$$

Therefore a_{li} must be a solution of $n(n+1)/2$ linear congruences. These congruences may have no solution, a unique solution, or a family of solutions. If at any stage it is impossible to solve the congruences (2), then there is no factorization A having the chosen main diagonal elements.

It is interesting to note that if $l=i$,

$$q_{(i-1)r+t,(i-1)r+j} \equiv 0 \bmod a_{ii}, \qquad t\leqq j.$$

The above method gives an algorithm for the complete solution of the equation $\sum_{m=0}^{s} A_m \cdot \times K_m X^m = 0$, if the matrices involved are square.

Results similar to those in this paper may be obtained, in some cases more readily, in the consideration of $\sum_{m=0}^{s} (X^m K_m) \cdot \times A_m = 0$, $\sum_{m=0}^{s} A_m \times \cdot (K_m X^m) = 0$, and $\sum_{m=0}^{s} (X^m K_m) \times \cdot A_m = 0$, and so on, where $A \times \cdot B = (a_{ij} B)$.

Bibliography

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

2. C. C. MacDuffee, *The theory of matrices*, Ergebnisse der Mathematik und ihrer Grenzgebeite, 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).

Michigan State College

AVERAGES OF CHARACTER SUMS

P. T. BATEMAN AND S. CHOWLA

Suppose that χ is a primitive residue character[1] modulo k, $k>1$, and that for y non-negative, $S(y) = \sum_{0 \leq l \leq y} \chi(l)$. It is important [see, for example, **11**] in the analytic theory of numbers to have as much information as possible about the sums $S(y)$, in particular about their maximum order of magnitude; it is known (cf. [**13**; **14**; **8**]), for example, that $S(y) < k^{1/2} \log k$, but unknown whether or not $M(\chi) = o(k^{1/2} \log k)$ as k tends to infinity, where $M(\chi)$ is the maximum of $|S(1)|, \cdots, |S(k-1)|$. Hua [**4**; **5**; **6**] has shown that it is often helpful to consider the averages $n^{-1} \sum_{m=0}^{n} S(m)$. In this paper we consider some further developments of this idea.

1. **Preliminaries.** We recall [**7**, pp. 483–486, 492–494] that if χ is a primitive residue character mod k and if $\tau(\chi) = \sum_{n=1}^{k} \chi(n) e^{2\pi i n/k}$, then $|\tau(\chi)| = k^{1/2}$ and

$$\sum_{n=1}^{k} \chi(n) e^{2\pi i m n/k} = \bar{\chi}(m) \tau(\chi) \tag{1}$$

for any integer m, $\bar{\chi}$ being the complex conjugate of χ.

Presented to the Society, April 30, 1949; received by the editors September 23, 1949.

[1] For the basic facts about residue characters see [**7**, pp. 401–414, 478–494]. Numbers in brackets refer to the bibliography.