

THE OHIO JOURNAL OF SCIENCE

Vol. 74

SEPTEMBER, 1974

No. 5

QUADRATIC MATRIX EQUATIONS¹

G. L. SHURBET, T. O. LEWIS, AND T. L. BOULLION

Department of Mathematics, Texas Tech University, Lubbock, Texas 79409

ABSTRACT

Necessary and sufficient conditions for the matrix equation $XAX=B$ and for the system of matrix equations $XAX=B$, $\sum_{j=1}^n C_{ij}XD_{ij}=E_i$, $i=1,2,\dots,m$, to be solvable are proven in this paper. If solvable, all solutions of the equation or system are determined. These results are used to obtain particular solutions of the general quadratic matrix equation $XAX+BX+XC=D$.

INTRODUCTION

Solutions of the matrix equation $XAX+BX+XC=D$ are important in many applications, e.g., see Potter (1966) who has solved a special case of the equation, but the general problem has not been solved. In this paper, additional particular solutions are obtained by the decomposition of D into a sum of three matrices. Unfortunately, there is no procedure for determining every permissible decomposition of D .

DEFINITIONS AND NOTATION

The letters A, B, \dots, R and X, Y will denote known and unknown matrices, respectively, of arbitrary orders up to the assumption that indicated operations are defined. All matrices are considered to be defined over the field of complex numbers; however, the results are valid over more general fields (Hurt and Waid, 1970). The Kronecker product of A and B , denoted $A \otimes B$, is defined by $A \otimes B = (a_{ij}B)$ where $A = (a_{ij})$. If B is a matrix, b will be used to represent B considered as a column vector with lexicographic order on the subscripts. Any solution of $AXA=A$ is called a generalized inverse of A , denoted A^- . For example, if $A=(1,2)$, then A^- will represent any column vector of the form $(1-2a,a)^T$ where a is arbitrary. Any square matrix A such that $A^2=B$ will be denoted $B^{1/2}$. For example $0^{1/2}$ will represent any matrix of the form $P^{-1}JP$ where P is an arbitrary invertible matrix and J contains zeros everywhere except for a one in the upper right position. Matrix J of order two also shows easily that not every matrix has a square root. The reader is referred to articles by Amir-Moez and Symrl and by Dade and Taussky, 1965, for known results about existence and determination of square roots of certain matrices.

PRELIMINARY RESULTS

Two preliminary theorems are proven in this section since they are essential to the main results.

Theorem 1. The matrix equation

$$(1) \quad XAX=B$$

is solvable if and only if $(AB)^{1/2}$ exists and

$$(2) \quad AA^-(AB)^{1/2}=(AB)^{1/2},$$

$$(3) \quad B(AB)^{1/2}-(AB)^{1/2}=B,$$

in which case every solution is of the form

¹Manuscript received December 18, 1972 (72-59).

$$(4) \quad X = A^-(AB)^{1/2} + (I - A^-A)\{B(AB)^{1/2} + U[I - (AB)^{1/2}(AB)^{1/2}]\},$$

where U is arbitrary.

Proof. Clearly (1) is equivalent to the system $AX = Y$, $XY = B$. For each fixed Y , this system is solvable (Morris and Odell, 1968) if and only if $AA^-Y = Y$, $BY^-Y = B$, and $Y^2 = AB$. Thus (1) is solvable if and only if $(AB)^{1/2}$ exists and (2), (3). Now it is known (Morris and Odell, 1968) that if the system is solvable, then every solution has the form $X = A^-Y + BY^- - A^-ABY^- + (I - A^-A)U(I - YY^-)$ where U is arbitrary. Then (4) follows immediately.

Theorem 2. The system of matrix equations

$$(5) \quad XAX = B,$$

$$(6) \quad \sum_{j=1}^n C_{ij}XD_{ij} = E_i, \quad i = 1, 2, \dots, m,$$

is solvable if and only if $(AB)^{1/2}$ exists and

$$(7) \quad AA^-(AB)^{1/2} = (AB)^{1/2},$$

$$(8) \quad B(AB)^{1/2} - (AB)^{1/2} = B,$$

$$(9) \quad F_i F_i^T G_i = G_i,$$

where

$$C_{ij}[A^-(AB)^{1/2} + (I - A^-A)B(AB)^{1/2}]D_{ij} = L_{ij},$$

$$C_{ij}(I - A^-A) = M_{ij},$$

$$[I - (AB)^{1/2}(AB)^{1/2}]D_{ij} = N_{ij},$$

$$P_i = E_i - \sum_{j=1}^n L_{ij},$$

$$Q_{ij} = M_{ij} \oplus N_{ij}^T,$$

$$R_i = \sum_{j=1}^n Q_{ij},$$

$$F_1 = R_1,$$

$$G_1 = p_1,$$

$$H_1 = R_1 p_1,$$

$$J_1 = I - R_1 R_1,$$

and for $i = 2, \dots, m$,

$$F_i = R_i J_{i-1},$$

$$G_i = p_i - R_i H_{i-1},$$

$$H_i = H_{i-1} + J_{i-1} F_i^T G_i,$$

$$J_i = J_{i-1}(I - F_i F_i^T),$$

in which case every solution is of the form

$$(10) \quad X = A^-(AB)^{1/2} + (I - A^-A)\{B(AB)^{1/2} + U_1[I - (AB)^{1/2}(AB)^{1/2}]\},$$

where U_1 is determined by

$$(11) \quad u_1 = H_m + J_m U,$$

and U is arbitrary.

Proof. By Theorem 1, (5) is solvable if and only if (7) and (8), in which case every solution has the form (10) where U_1 is arbitrary. Then the system

(5), (6) is solvable if and only if (6) is solvable for U_1 when (10) is substituted for X in (6). This substitution yields

$$R_i u_1 = p_i, \quad i = 1, 2, \dots, m.$$

But this system is solvable (Morris and Odell, 1968) if and only if (9), in which case every solution has the form (11).

A special case of Theorem 2 is presented next, with notation simplified considerably, since it is the basis for the main results of this paper.

Corollary 1. The system of matrix equations

$$\begin{aligned} XAX &= D_1, \\ BX &= D_2, \\ XC &= D_3, \end{aligned}$$

is solvable if and only if $(AD_1)^{1/2}$ exists and

$$\begin{aligned} AA^-(AD_1)^{1/2} &= (AD_1)^{1/2}, \\ D_1(AD_1)^{1/2} - (AD_1)^{1/2} &= D_1, \\ BEB^-GF &= G, \\ KK^-j &= j, \end{aligned}$$

where

$$\begin{aligned} E &= I - A^-A, \\ F &= I - (AD_1)^{1/2}(AD_1)^{1/2-}, \\ G &= D_2 - BA^-(AD_1)^{1/2} - BED_1(AD_1)^{1/2-}, \\ H &= -EB^-BE, \\ J &= D_3 - A^-(AD_1)^{1/2}C - ED_1(AD_1)^{1/2-}C - EB^-GFC, \\ K &= (E+H) \oplus (FC)^T, \end{aligned}$$

in which case every solution has the form

$$X = A^-(AD_1)^{1/2} + ED_1(AD_1)^{1/2-} + EB^-GF + E(I - B^-BE)U_1F,$$

where U_1 is determined by

$$u_1 = K^-j + (I - K^-K)U,$$

and U is arbitrary.

MAIN RESULTS

Consider the matrix equation

$$(12) \quad XAX + BX + XC = D.$$

Particular solutions of this equation can be obtained by writing $D = D_1 + D_2 + D_3$ where D_1, D_2, D_3 satisfy the conditions of Corollary 1. Unfortunately, the conditions do not give a method for determining every decomposition of D such that the matrix equation is solvable. Six obvious specializations of the corollary which may give particular solutions are now listed. Case (vi) is given in Theorem 3 as an example.

- (i) $D_1 = D_2 = D_3 = D = 0$,
- (ii) $D_1 = D_2 = 0, D_3 = D$,
- (iii) $D_1 = 0, D_2 = D, D_3 = 0$
- (iv) $D_1 = D, D_2 = D_3 = 0$,
- (v) $D_1 = D, D_2 + D_3 = 0$,
- (vi) $D_1 = 0, D_2 + D_3 = D$.

Theorem 3. If $AA^{-1} = 0$ and $EE^{-1} = I$ where

$$E = [B(I - A^{-1}A)] \oplus (I - 0^{-1}0^{-1})^T + (I - A^{-1}A) \oplus [(I - 0^{-1}0^{-1})C]^T$$

and $F = D - BA^{-1} - A^{-1}C$, then

$$X = A^{-1} + (I - A^{-1}A)U_1(I - 0^{-1}0^{-1}),$$

where U is determined by

$$u_1 = E^{-1}f + (I - E^{-1}E)U,$$

and U is arbitrary, is a solution of (12).

Theorems 4-7 yield particular solutions of (12) obtained simply by factoring. However, the theorems are useful in conjunction with Theorems 8-9. Proof of Theorem 4 is given to show the method.

Theorem 4. If A is square, $BA = B$, and $BC = -D$, then $X = -B$ is a solution of (12). If in addition $AA^{-1}C = C$, then $X = -A^{-1}C + (I - A^{-1}A)U$, where U is arbitrary, is a solution of (12).

Proof. The first conditions permit the factorization $XAX + BX + XC - D = (X + B)(AX + C)$. Then obviously $X = -B$ is a solution of (12). Further, if $AA^{-1}C = C$ then $AX = -C$ is solvable (Penrose, 1955) and every solution has the form $X = -A^{-1}C + (I - A^{-1}A)U$, where U is arbitrary. Again this X is clearly a solution of (12).

Theorem 5. If $DA = B$ and $DC = -D$, then $X = -D$ is a solution of (12). If in addition $AA^{-1}C = C$, then $X = -A^{-1}C + (I - A^{-1}A)U$, where U is arbitrary, is a solution of (12).

Theorem 6. If A is square, $AC = C$, and $BC = -D$, then $X = -C$ is a solution of (12). If in addition $BA^{-1}A = B$, then $X = -BA^{-1} + U(I - AA^{-1})$, where U is arbitrary, is a solution of (12).

Theorem 7. If $AD = C$ and $BD = -D$, then $X = -D$ is a solution of (12). If in addition $BA^{-1}A = B$, then $X = -BA^{-1} + U(I - AA^{-1})$, where U is arbitrary, is a solution of (12).

It is easily seen from (12) that both B and C must be square matrices and thus may be invertible. Theorems 8-11 utilize this possibility.

Theorem 8. Suppose $\det B \neq 0$. Then (12) is solvable if and only if

$$(13) \quad YAB^{-1}Y + BY + YC = BD$$

is solvable, in which case $X = B^{-1}Y$ is a solution of (12).

Proof. Suppose (12) is solvable. Then there exists D_2 such that $BX = D_2$ or $X = B^{-1}D_2$ and X satisfies (12). Then $B^{-1}D_2AB^{-1}D_2 + D_2 + B^{-1}D_2C = D$ and $D_2AB^{-1}D_2 + BD_2 + D_2C = BD$, i.e., (13) is solvable. Now suppose (13) is solvable. Then $B^{-1}YAB^{-1}Y + BB^{-1}Y + B^{-1}YC = D$, i.e., $X = B^{-1}Y$ is a solution of (12).

Theorem 9. Suppose $\det C \neq 0$. Then (12) is solvable if and only if

$$(14) \quad YC^{-1}AY + BY + YC = DC$$

is solvable, in which case $X = YC^{-1}$ is a solution of (12).

The proof, similar to that of Theorem 8, is omitted.

Clearly, methods indicated previously may be used to find particular solutions of (13) or (14) which then give solutions to (12). Theorems 10-11 are given as examples.

Theorem 10. Suppose $\det B \neq 0$. If $C = -AB^{-1}D$, then $X = B^{-1}D$ is a solution of (12). If in addition $A^{-1}A = I$, then $X = -BA^{-1} + B^{-1}U(I - AA^{-1})$ is a solution of (12) for each arbitrary matrix U .

Proof. Proof by substitution is very simple; however, a longer proof is given to show how other solutions may be obtained. By Theorem 6, (12) is solvable

and $X = B^{-1}Y$ if and only if $YAB^{-1}Y + BY + YC - BD = 0$ is solvable. Now if $C = -AB^{-1}D$, then

$$\begin{aligned} YAB^{-1}Y + BY + YC - BD &= YAB^{-1}Y + BY - YAB^{-1}D - BD \\ &= (YAB^{-1} + B)(Y - D), \end{aligned}$$

thus $X = B^{-1}D$ is a solution of (12). Also if $YAB^{-1} = -B$, i.e., if $A^{-1}A = I$ so that $Y = -B^2A^{-1} + U(I - AA^{-1})$, then $X = B^{-1}Y = -BA^{-1} + B^{-1}U(I - AA^{-1})$, where U is arbitrary, is a solution of (12). Clearly, other factorizations may be useful in determining other solutions of (12).

Theorem 11. Suppose $\det C \neq 0$. If $B = -DC^{-1}A$, then $X = DC^{-1}$ is a solution of (12). If in addition $AA^{-1} = I$, then $X = -A^{-1}C + (I - A^{-1}A)UC^{-1}$ is a solution of (12) for each arbitrary matrix U .

The proof is similar to that of Theorem 10.

REFERENCES CITED

- Dade, E. C., and O. Taussky.** 1965. Some new results connected with matrices of rational integers. *Proc. Symp. Pure Math.* 8: 78-88.
- Amir-Moez, A. R., and S. Smyrl.** (In press). Square roots of a positive matrix and polar decompositions. *J. Indust. Math.*
- Hurt, M. J., and C. Waid.** 1970. A generalized inverse which gives all the integral solutions to a system of linear equations. *SIAM J. Appl. Math.* 19: 547-550.
- Morris, G. L., and P. L. Odell.** 1968. Common solutions for n matrix equations with applications. *J. Assoc. Comput. Mach.*, 15: 272-274.
- Penrose, R.** 1955. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* 51: 406-413.
- Potter, J. E.** 1966. Matrix quadratic solutions. *SIAM J. Appl. Math.* 14: 496-501.