SOLUTION OF CERTAIN MATRIX EQUATIONS

JOHN JONES, JR.

ABSTRACT. The main purpose of this paper is to obtain solutions of matrix equations of the following types, $AX-XB=C$, $XDX+AX+XB+C=0$, in which case X is an unknown n by n matrix and A, B, C, D are n by n matrices having elements belonging to the field C of complex numbers. Results obtained extend those of W. E. Roth, J. E. Potter and others concerning the existence and the representation of solutions X of the above equations.

1. **Introduction.** The main purpose of this paper is to obtain solutions of matrix equations of the following types

$$(1.1) \qquad\qquad AX - XB = C,$$

$$(1.2) \qquad XDX + AX + XB + C = 0,$$

in which X is an unknown n by n matrix having elements belonging to the field C of complex numbers. The matrices A, B, C, D are n by n matrices having elements belonging to C. The elements of all matrices and the coefficients of all polynomials used throughout this paper will belong to the field C and so the similarity of matrices will be valid under the rational operations of C. Capital letters will denote matrices and I will be the n by n identity matrix. The generalized inverse of a matrix as defined by R. Penrose [1] will be denoted by †. We will make use of the notation used by W. E. Roth [3], [4] throughout this paper.

Equations of the types (1.1), (1.2) are of considerable interest due to their numerous applications in the areas of ordinary and partial differential equations, difference equations, optimal control theory, structural analysis, stability theory, filtering theory and elsewhere.

2. **Solutions of equation (1.1).** In this section solutions of (1.1) will be obtained in terms of generalized inverses of certain matrices. Let the $2n$ by $2n$ matrices $R, \tilde{R}$ be denoted as in W. E. Roth [3] by

$$(2.1) \qquad R = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \qquad \tilde{R} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

---

where A, B, C are given in (1.1). A sufficient, though not necessary condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case the solution of (1.1) not only exists, but is unique. $f_A(\lambda) = |A - \lambda I|$, $f_B(\lambda) = |B - \lambda I|$ are the characteristic polynomials of A, B respectively. W. E. Roth [3] established the following result.

THEOREM 2.1 (ROTH [3]). *The necessary and sufficient condition that equation (1.1), where A, B, C are square matrices of order n with elements in C, have a solution X with elements in C, is that the matrices given in (2.1) be similar.*

We will next obtain the form of a solution X of (1.1) under the hypotheses of Theorem 2.1 above. Let the following $2n$ by $2n$ matrices be denoted by

$$(2.2) \quad f_A(R) = \begin{pmatrix} 0 & M \\ 0 & U \end{pmatrix}, \quad f_B(R) = \begin{pmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

where the n by n matrices U, M, $\hat{M}$, $\hat{N}$ are polynomials in the matrices A, B, C given in (1.1) and $f_A(\lambda)$, $f_B(\lambda)$ are the characteristic polynomials of A, B respectively. We will show that the hypotheses of Theorem 2.1 above imply that there exists a solution X of (1.1) of the form

$$(2.3) \quad X = \hat{N}^\dagger \hat{M} - MU^\dagger + \hat{N}^\dagger \hat{N} M U^\dagger.$$

THEOREM 2.2. *If the matrices R, $\tilde{R}$ of (2.1) are similar then equation (1.1) has a solution X of the form (2.3) above.*

PROOF. From Theorem 2.1 above let X be a solution of (1.1), then we have the following equation,

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & M \\ 0 & U \end{pmatrix} \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}$$

$$(2.4) \qquad = \begin{pmatrix} 0 & M + XU \\ 0 & U \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} f_A(R) \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} f_A(A) & 0 \\ 0 & f_A(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & f_A(B) \end{pmatrix}.$$

Equation (2.4) implies that $M + XU = 0$. Similarly we have

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} \hat{N} & \hat{M} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}$$

$$(2.5) \qquad = \begin{pmatrix} \hat{N} & -\hat{N}X + \hat{M} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} f_B(R) \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} f_B(A) & 0 \\ 0 & f_B(B) \end{pmatrix} = \begin{pmatrix} f_B(A) & 0 \\ 0 & 0 \end{pmatrix}.$$

Equation (2.5) implies that $\hat{M}-\hat{N}X=0$, and thus a solution X of (1.1) must satisfy the pair of equations

$$(2.6) \qquad M + XU = 0, \qquad \hat{M} - \hat{N}X = 0.$$

A necessary and sufficient condition for $\hat{N}X=\hat{M}$ to have a solution is that $\hat{N}\hat{N}^\dagger\hat{M}=\hat{M}$, and the general solution X_1 is given by

$$(2.7) \qquad X_1 = \hat{N}^\dagger\hat{M} + Y_1 - \hat{N}^\dagger\hat{N}Y_1, \qquad Y_1 \text{ arbitrary.}$$

A necessary and sufficient condition for $XU=-M$ to have a solution is that $MU^\dagger U=M$, and the general solution X_2 is given by

$$(2.8) \qquad X_2 = -MU^\dagger + Y_2 - Y_2UU^\dagger, \qquad Y_2 \text{ arbitrary.}$$

A necessary and sufficient condition for the pair of equations given in (2.6) to have a common solution is that each equation individually have a solution and $-\hat{N}M=\hat{M}U$. A common solution X is given by (2.3) above. Thus a solution $X=X_1=X_2$ of (1.1) is of the form (2.3) where $Y_2=\hat{N}^\dagger\hat{M}$ and $Y_1=-MU^\dagger$.

We next establish other sufficient conditions that equation (1.1) has a solution X with elements in C.

THEOREM 2.3. *Let $f_\alpha(\lambda)$ be a polynomial of degree $n\geq 1$ in λ with coefficients in C such that*

$$(2.9) \qquad f_\alpha(R) = \begin{pmatrix} V & N \\ 0 & 0 \end{pmatrix}, \qquad R = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

where V, N are polynomials in A, B, C and V^{-1} exists. Then a solution X of $N-VX=0$ is also a solution of (1.1).

PROOF. As mentioned earlier a sufficient, though not necessary condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case a solution of (1.1) not only exists, but is unique. But this is also necessary for the hypothesis here implies that $f_\alpha(B)=0$, whence $f_\alpha(\lambda)$ is a multiple of the minimum polynomial of B, and that $V=f_\alpha(A)$ is nonsingular, whence $f_\alpha(\lambda)$ cannot vanish for any characteristic root of A.

The matrices $f_\alpha(R)$ and R commute which implies the following identities:

$$(2.10) \qquad AV = VA, \qquad AN - VC - NB = 0.$$

Let X be a solution of $N-VX=0$, then using (2.10) we have the following equation

$$(2.11) \quad 0 = A(N - VX) = VC + NB - AVX = V[C + XB - AX]$$

and since V^{-1} exists by hypothesis, X is a solution of (1.1).

THEOREM 2.4. *Let $f_\beta(\lambda)$ be a polynomial of degree $n \geq 1$ in λ with coefficients in C such that*

$$(2.12) \qquad f_\beta(R) = \begin{pmatrix} 0 & N \\ 0 & M \end{pmatrix}, \qquad R = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

where N, M are polynomials in A, B, C and M^{-1} exists. Then a solution X of $N + XM = 0$ is also a solution of (1.1).

PROOF. Similarly as in the proof of Theorem 2.3 above, a sufficient, though not necessary, condition that the matrices R and $\tilde{R}$ be similar is, of course, that A and B have no common characteristic root. In this case a solution of (1.1) not only exists, but is unique. But this is also necessary for the hypothesis here implies that $f_\beta(A) = 0$, whence $f_\beta(\lambda)$ is a multiple of the minimum polynomial of A and that $M = f_\beta(B)$ is nonsingular, whence $f_\beta(\lambda)$ cannot vanish for any characteristic root of B.

The matrices of $f_\beta(R)$ and R commute which implies the following identities:

$$(2.13) \qquad BM = MB, \qquad AN + CM - NB = 0.$$

Let X be a solution of $N + XM = 0$, then using (2.13) we have the following equation

$$(2.14) \quad \begin{aligned} 0 &= -(N + XM)B = -NB - XMB = -CM - AN - XMB \\ &= -CM + AXM - XBM = (-C + AX - XB)M \end{aligned}$$

and since M^{-1} exists by hypothesis X is a solution of (1.1).

3. **Solutions of equation (1.2).** For numerous applications of solutions of (1.2) see J. E. Potter [2]. The purpose of this section is to extend some of the results of J. E. Potter [2] and W. E. Roth [4].

Let $R, f_\gamma(R)$ be the following $2n$ by $2n$ matrices formed from the matrices A, B, C, D of (1.2),

$$(3.1) \qquad R = \begin{pmatrix} -B & D \\ -C & A \end{pmatrix}, \qquad f_\gamma(R) = \begin{pmatrix} U & M \\ V & N \end{pmatrix},$$

where $f_\gamma(\lambda)$ is any polynomial of degree $n \geq 1$ in λ having coefficients in C. U, V, M, N are polynomials in A, B, C, D. For the case $D = I$ the following result was established.

THEOREM 3.1 (W. E. ROTH [4]). *If $f_\gamma(\lambda)$ is a polynomial of degree $\gamma \leq n$ with coefficients in C such that $f_\gamma(R)$ is given by (3.1) and is of rank n and if M is nonsingular, then X such that $(X, I)f_\gamma(R) = (0, 0)$ is a solution of (1.2) with elements in C and $X_1 = -M^{-1}(X + B)M - A$ is likewise a solution.*

We next establish the following extension of Theorem 3.1 above.

THEOREM 3.2.   *Let $f_\gamma(\lambda)$ be any polynomial of degree $n \geq 1$ in λ with coefficients belonging to C such that $R, f_\gamma(R)$ are given by (3.1). Then a solution of $(X, I)f_\gamma(R) = (0, 0)$ with U^{-1} or M^{-1} existing, or a solution of $f_\gamma(R)\binom{I}{-X} = \binom{0}{0}$ with M^{-1} or N^{-1} existing is also a solution of (1.2).*

PROOF.   We will first show that if a solution of $(X, I)f_\gamma(R) = (0, 0)$ exists with U^{-1} or M^{-1} existing then such a solution is also a solution of (1.2). Let such a solution X exist. Now the matrices $R, f_\gamma(R)$ commute so we have the following matrix identities:

$$\text{(3.2)} \quad \begin{aligned} DV + UB &= BU - MC, & UD + MA &= DN - BM, \\ AV + VB &= CU - NC, & CM + VD &= -NA + AN. \end{aligned}$$

Making use of (3.2) and the above solution X we have

$$\begin{aligned} 0 &= -(XM + N)C \\ &= -XUB - XMC + XUB - NC \\ &= -X(UB + MC) + XUB - NC \\ &= -X(BU - DV) + XUB - NC \\ \text{(3.3)} \quad &= -XBU - XDXU + XUB - NC \\ &= -XDXU - AXU - XBU + AXU + XUB - NC \\ &= -XDXU - AXU - XBU - AV - VB - NC \\ &= -XDXU - XBU - AXU - CU \\ &= -(XDX + AX + XB + C)U. \end{aligned}$$

U^{-1} exists by hypothesis and (3.3) implies that X is a solution of (1.2). Next let M^{-1} exist, then using (3.2) we have

$$\begin{aligned} 0 &= (XM + N)A \\ &= -VD + XMA + NA + VD \\ &= XUD + XMA + NA + VD \\ &= X(DN - BM) + NA + VD \\ \text{(3.4)} \quad &= XDN - AXM - XBM + AXM + NA - XUD \\ &= -XDXM - AXM - XBM + NA - AN + VD \\ &= -XDXM - AXM - XBM - CM \\ &= -(XDX + AX + XB + C)M. \end{aligned}$$

Equation (3.4) implies that X is a solution of (1.2).
    Next let X be a solution of $f_\gamma(R)\binom{I}{-X} = \binom{0}{0}$ with N^{-1} or M^{-1} existing.

Making use of (3.2) we have

$$
\begin{aligned}
0 &= D(V - NX) \\
&= (BU - UB - MC) - (UD + MA + BM)X \\
&= (BMX - MXB - MC) - (MXD + MA + BM)X \\
&= -M(XDX + AX + XB + C)
\end{aligned}
$$

(3.5)

which implies that X is a solution of (1.2). Similarly we have

$$
\begin{aligned}
0 &= -A(V - NX) \\
&= (-CU + NC + VB) + (CM + VD + NA)X \\
&= -CMX + NC + VB + CMX + NXDX + NAX \\
&= -AV + CU + NXDX + NAX \\
&= VB + NC + NXDX + NAX \\
&= NXB + NC + NXDX + NAX \\
&= N(XDX + AX + XB + C).
\end{aligned}
$$

(3.6)

Equation (3.6) implies that X is a solution of (1.2).

We next establish other sufficient conditions for the existence of solutions of (1.2).

THEOREM 3.3. *Let R, $f_\alpha(R)$ be given as in (3.1) where M^{-1} exists and $V = VU^\dagger U$, $VU^\dagger = NM^{-1}$. Then $X = -NM^{-1}$ is a solution of (1.2).*

PROOF. Now $-N = -NM^{-1}M = XM$ and so $XM = -N$ has a solution. Also $XU = -NM^{-1}U = -VU^\dagger U = -V$. Therefore $(X, I)f_\alpha(R) = (0, 0)$ has a solution and since M^{-1} exists by Theorem 3.2, $X = -NM^{-1}$ is a solution of (1.2).

THEOREM 3.4. *Let R, $f_\alpha(R)$ be given as in (3.1) where M^{-1} exists and $V = NN^\dagger V$, $U = MN^\dagger V$ hold. Then $X = M^{-1}U$ is a solution of (1.2).*

PROOF. $V = NN^\dagger V$ is a necessary and sufficient condition that $NX = V$ have a solution. Also

$$
U = MM^{-1}U = MX
$$

and

$$
NX = NM^{-1}U = NM^{-1}(MN^\dagger V) = NN^\dagger V = V.
$$

Therefore $f_\alpha(R)\binom{I}{-X} = \binom{0}{0}$, has a solution and since M^{-1} exists, by Theorem 3.2, $X = M^{-1}U$ is a solution of (1.2).

REFERENCES

1. R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc. **51** (1955), 406–413. MR **16,** 1082.

2. J. E. Potter, *Matrix quadratic solutions*, SIAM J. Appl. Math. **14** (1966), 496–501. MR **34** #1341.

3. W. E. Roth, *The equations $AX - YB = C$, and $AX - XB = C$ in matrices*, Proc. Amer. Math. Soc. **3** (1952), 392–396. MR **13,** 900.

4. ———, *On the matrix equation $X^2 + AX + XB + C = 0$*, Proc. Amer. Math. Soc. **1** (1950), 586–589. MR **12,** 471.

DEPARTMENT OF MATHEMATICS, GEORGE WASHINGTON UNIVERSITY, WASHINGTON, D.C. 20006

DEPARTMENT OF MATHEMATICS, AIR FORCE INSTITUTE OF TECHNOLOGY, DAYTON, OHIO 45433