SIAM J. APPL. MATH.
Vol. 31, No. 4, December 1976

HERMITIAN AND NONNEGATIVE DEFINITE SOLUTIONS OF LINEAR MATRIX EQUATIONS*

C. G. KHATRI AND SUJIT KUMAR MITRA†

Abstract. Necessary and sufficient conditions for existence and expressions for general Hermitian and nonnegative definite solutions are obtained for the following three systems of linear equations: (I) $AX = C$, (II) $AX = C$, $XB = D$, (III) $AXB = C$.

1. Introduction. Algebraic solutions of linear matrix equations using generalized inverses of matrices have a history dating back to Bjerhammer [3]. Some other contributors in this area are Penrose [8], [9], Rao [10], Morris and Odell [7], Rao and Mitra [12] and Mitra [6]. For some applications, however, the only solutions that are relevant may be those where the solution matrix is Hermitian or nonnegative definite. The MINQUE estimate of covariance components in a covariance components model given by Rao [11], for example, requires the solution of a matrix equation, linear in Σ (the covariance matrix of the model), where it is obviously desirable to restrict the solution Σ to be nonnegative definite. Similar problems also occur in load slow analysis and short circuit studies in power systems (see, e.g., Stagg and El-Abiad [13]).

Keeping such applications in view, the authors in the present paper study the following three systems of linear equations: (I) $AX = C$, (II) $AX = C$, $XB = D$ and (III) $AXB = C$, obtain necessary and sufficient conditions for the existence of a Hermitian or nonnegative definite solution, and also the general class of such solutions, for each case separately. This paper thus partially fills a gap between the well-known results on generalized inverse solutions of linear matrix equations (see references cited earlier and also the recent survey in Ben-Israel and Greville [2]) and the rich literature on positive definite solutions of linear matrix equations (such as Lyapunov type stability theorems, see, e.g., Hill [4], Ben-Israel and Berman [1]).

We consider matrices over the complex field. $\mathscr{C}^{m\times n}$ represents the linear space of all complex matrices of order $m \times n$. Matrices are denoted by capital letters such as A, B, X etc. For a matrix A, $\mathscr{M}(A)$ denotes its column span, A^* its complex conjugate transpose, a generalized inverse of A denoted by A^- is a matrix which satisfies the equation $AA^-A = A$ (Rao and Mitra [12]).

2. The main results.

THEOREM 2.1. *Let A and C be given matrices in $\mathscr{C}^{m\times n}$ such that the equation*

$$AX = C \tag{2.1}$$

is consistent. The equation (2.1) *has a Hermitian solution if and only if*

$$CA^* \quad \text{is Hermitian,} \tag{2.2}$$

* Received by the authors February 3, 1975, and in revised form September 16, 1975.

† Indian Statistical Institute, New Delhi 29, India. The first author is on leave from Gujarat University.

in which case a general Hermitian solution is

$$A^-C+C^*(A^-)^*-A^-AC^*(A^-)^*+(I-A^-A)U(I-A^-A)^*, \tag{2.3}$$

where A^- is an arbitrary g-inverse of A and U an arbitrary Hermitian matrix in $\mathscr{C}^{n\times n}$.

Proof. Let X be an Hermitian solution of (2.1). Then $CA^*=AXA^*$ is clearly Hermitian. This shows the necessity of (2.2). For sufficiency, check that when CA^* is Hermitian so are AC^* and $X_0=A^-C+C^*(A^-)^*-A^-AC^*(A^-)^*$ and that X_0 satisfies (2.1).

A general Hermitian solution is obtained by adding to X_0 a general Hermitian solution of the homogeneous equation $AX=0$. Solution (2.3) therefore follows from Corollary 1 to Lemma 2.3.1 in Rao and Mitra [12, p. 26].

THEOREM 2.2. *The equation* (2.1) *has a nonnegative definite solution if and only if*

$$CA^* \quad \text{is nonnegative definite,} \qquad \operatorname{rank} CA^*=\operatorname{rank} C, \tag{2.4}$$

in which case a general nonnegative definite solution is

$$C^*(CA^*)^-C+(I-A^-A)U(I-A^-A)^*, \tag{2.5}$$

where $(CA^*)^-$ and A^- are arbitrary g-inverses of CA^* and A respectively and U is an arbitrary nonnegative definite matrix in $\mathscr{C}^{n\times n}$.

Proof. Let X be a nonnegative definite solution of (2.1). Then $CA^*=AXA^*$ is clearly nonnegative definite and $\operatorname{rank} CA^*=\operatorname{rank} AXA^*=\operatorname{rank} AX=\operatorname{rank} C$. This shows the necessity of (2.4). For sufficiency, check that when (2.4) is true, $\mathscr{M}(C)=\mathscr{M}(CA^*)$. Hence by Lemma 2.2.4 (iii) of Rao and Mitra [12, p. 21], $C^*(CA^*)^-C$ is invariant under the choice of a g-inverse of CA^*. Since a nonnegative definite matrix such as CA^* has a nonnegative definite g-inverse, it is seen that $X_0=C^*(CA^*)^-C$ is nonnegative definite. Also,

$$AX_0=AC^*(CA^*)^-C=CA^*(CA^*)^-C=C.$$

Sufficiency of (2.4) is thus established.

Since X_0 satisfies (2.1) and $\operatorname{rank} X_0=\operatorname{rank} C$, by Note 1 following Lemma 2.2 in Mitra [5], it is seen that $X_0=GC=GAX_0$ for some g-inverse G of A. As in Theorem 2.1, a general Hermitian solution to (2.1) is given by

$$X=X_0+V,$$

where $V=(I-GA)U(I-GA)^*$ and U is Hermitian. If U is nonnegative definite, then so are V and X. Also since

$$\begin{aligned}(I-GA)X(I-GA)^*&=(I-GA)X_0(I-GA)^*+(I-GA)V(I-GA)^*\\&=(I-GA)V(I-GA)^*=V,\end{aligned}$$

if X is nonnegative definite, so is V. This shows that each nonnegative definite solution X of (2.1) can be expressed as

$$X=X_0+(I-GA)U(I-GA)^* \tag{2.6}$$

for a proper choice of the nonnegative definite matrix U. Let A^- be an arbitrary

g-inverse of A. Since $(I-A^-A)(I-GA)=(I-GA)$, substitution of $(I-A^-A)(I-GA)$ for $(I-GA)$ in (2.6) shows that (2.5) indeed provides a general nonnegative definite solution to (2.1).

THEOREM 2.3. *Let A and C be given matrices in $\mathscr{C}^{m\times n}$ and B, D be given matrices in $\mathscr{C}^{n\times p}$ such that the equations*

$$AX=C,\quad XB=D \tag{2.7}$$

are jointly consistent.

(a) *These equations have a common Hermitian solution if and only if*

$$M=\begin{pmatrix} CA^* & CB \\ D^*A^* & D^*B\end{pmatrix} \tag{2.8}$$

is Hermitian, in which case a general Hermitian solution is

$$\begin{pmatrix} A \\ B^*\end{pmatrix}^-\begin{pmatrix} C \\ D^*\end{pmatrix}+(C^*:D)\left[\left(\begin{pmatrix} A \\ B^*\end{pmatrix}^-\right)^*-\begin{pmatrix} A \\ B^*\end{pmatrix}^- M\left[\begin{pmatrix} A \\ B^*\end{pmatrix}^-\right]^*\right.$$

$$+\left[I-\begin{pmatrix} A \\ B^*\end{pmatrix}^-\begin{pmatrix} A \\ B^*\end{pmatrix}\right]U\left[I-\begin{pmatrix} A \\ B^*\end{pmatrix}^-\begin{pmatrix} A \\ B^*\end{pmatrix}\right]^*,$$

where U is an arbitrary Hermitian matrix in $\mathscr{C}^{n\times n}$.

(b) *These equations have a common nonnegative definite solution if and only if*

$$M \text{ is nonnegative definite} \quad \text{and} \quad \operatorname{rank} M=\operatorname{rank}(C^*:D), \tag{2.9}$$

in which case a general nonnegative definite solution is

$$(C^*:D)M^-\begin{pmatrix} C \\ D^*\end{pmatrix}+\left[I-\begin{pmatrix} A \\ B^*\end{pmatrix}^-\begin{pmatrix} A \\ B^*\end{pmatrix}\right]U\left[I-\begin{pmatrix} A \\ B^*\end{pmatrix}^-\begin{pmatrix} A \\ B^*\end{pmatrix}\right]^* \tag{2.10}$$

where U is an arbitrary nonnegative definite matrix in $\mathscr{C}^{n\times n}$.

Proof. Observe that for $AX=C$, $XB=D$ to have a common Hermitian (nonnegative definite) solution it is necessary and sufficient that the equation

$$\begin{pmatrix} A \\ B^*\end{pmatrix}X=\begin{pmatrix} C \\ D^*\end{pmatrix}$$

has a Hermitian (nonnegative definite) solution. Theorem 2.3 therefore follows from Theorems 2.1 and 2.2.

In the next two theorems we look for Hermitian and nonnegative definite solutions of the consistent matrix equation $AXB=C$. As in Mitra [6], we assume here that the coefficient matrices A and B are both nonnegative definite. This involves no loss of generality since whenever the equation $AXB=C$ is consistent,

$$AXB=C \Leftrightarrow A^*AXBB^*=A^*CB^*.$$

THEOREM 2.4. *Let A and B be given nonnegative definite matrices in $\mathscr{C}^{n\times n}$ and C be a given matrix in $\mathscr{C}^{n\times n}$ such that the equation*

$$AXB=C \tag{2.11}$$

is consistent. (2.11) *has a Hermitian solution if and only if*

$$T = B(A+B)^{-}C(A+B)^{-}A \tag{2.12}$$

is Hermitian, in which case a general Hermitian solution is

$$\begin{aligned} X = {} & (A+B)^{-}(C+C^{*}+Y+Z)[(A+B)^{-}]^{*}+U \\ & -(A+B)^{-}(A+B)U(A+B)[(A+B)^{-}]^{*}, \end{aligned} \tag{2.13}$$

where $(A+B)^{-}$ *is a g-inverse of* $(A+B)$, U *is an arbitrary Hermitian matrix in* $\mathscr{C}^{n\times n}$ *and* Y, Z *are arbitrary Hermitian solutions of the equations*

$$\begin{aligned} Y(A+B)^{-}B &= C(A+B)^{-}A, \\ A(A+B)^{-}Z &= B(A+B)^{-}C \end{aligned} \tag{2.14}$$

obtained as in Theorem 2.1.

Proof. We first show that for the equation $AXB = C$ to have a Hermitian solution, it is necessary and sufficient that the equations $AXB = C$ and $BXA = C^{*}$ have a common solution. If the Hermitian matrix X satisfies $AXB = C$, the equations $AXB = C$ and $BXA = C^{*}$ clearly have a common solution, in fact, a common Hermitian solution. Conversely, if X satisfies both $AXB = C$ and $BXA = C^{*}$, $(X+X^{*})/2$ is a Hermitian solution of $AXB = C$. Theorem 2.4 therefore follows from Theorem 2.1 of Mitra [6]. We leave it to the reader to verify the necessary details. It is a routine matter as in Mitra [6] to check that X as defined in (2.13) is a Hermitian solution of (2.11). Conversely, let $\hat{X}$ be a Hermitian solution of (2.11). Put $Y = A\hat{X}A$, $Z = B\hat{X}B$ and check that Y and Z are Hermitian solutions of (2.14). Also $X = \hat{X}$ satisfies the equation

$$(A+B)X(A+B) = C+C^{*}+Y+Z$$

of which $X_0 = (A+B)^{-}(C+C^{*}+Y+Z)[(A+B)^{-}]^{*}$ is another Hermitian solution.

This shows that X can always be represented as $X_0 + W$, where W is a Hermitian solution of the homogeneous equation $(A+B)X(A+B) = 0$. An application of Corollary 1 to Lemma 2.3.1 in Rao and Mitra [12, p. 36] establishes the expression (2.13) for X.

THEOREM 2.5. (2.11) *has a nonnegative definite solution if and only if* T *as defined in* (2.12) *is nonnegative definite and*

$$\operatorname{rank} T = \operatorname{rank}\{A(A+B)^{-}C^{*}\} = \operatorname{rank}\{B(A+B)^{-}C\}, \tag{2.15}$$

in which case a general nonnegative definite solution is given by

$$\begin{aligned} X = {} & (A+B)^{-}(C+C^{*}+Y+Z)[(A+B)^{-}]^{*} \\ & +[I-(A+B)^{-}(A+B)]U[I-(A+B)^{-}(A+B)]^{*}, \end{aligned} \tag{2.16}$$

where Y *and* Z *are arbitrary nonnegative definite solutions of* (2.14) *such that* $C+C^{*}+Y+Z$ *is nonnegative definite,* $(A+B)^{-}$ *is an arbitrary g-inverse of* $(A+B)$ *and* U *is an arbitrary nonnegative definite matrix in* $\mathscr{C}^{n\times n}$.

Proof. For the necessity part observe that when (2.11) is consistent,

$$T = B(A+B)^{-}AXB(A+B)^{-}A.$$

Further, for nonnegative definite matrices A and B, $B(A+B)^{-}A$ is Hermitian and, in fact, nonnegative definite (see Rao and Mitra [12, p. 191]). Hence when X is nonnegative definite so is T and

$$\begin{aligned}\operatorname{rank} T &= \operatorname{rank} B(A+B)^{-}AXB\\ &= \operatorname{rank} B(A+B)^{-}C\\ &= \operatorname{rank} AXB(A+B)^{-}A\\ &= \operatorname{rank} C(A+B)^{-}A.\end{aligned}$$

We now proceed to prove the sufficiency part and along with it the rest of the theorem. We first show that when these conditions are true, nonnegative definite solutions Y, Z to (2.14) can be so determined that

$$C+C^{*}+Y+Z$$

is nonnegative definite. To simplify writing, let us rewrite (2.14) as $YE=F$, $KZ=L$, where E, F, K and L have obvious interpretations. Check that $LK^{*}=T=E^{*}F$. Also (2.15) implies rank $T=$ rank $F=$ rank L. Hence if T is nonnegative definite, by Theorem 2.2 general nonnegative definite solutions to (2.14) can be written as

$$\begin{aligned}Y&=FT^{-}F^{*}+(I-EE^{-})^{*}V(I-EE^{-}),\\ Z&=L^{*}T^{-}L+(I-K^{-}K)W(I-K^{-}K)^{*}.\end{aligned}$$

Hence

$$\begin{aligned}C+C^{*}+Y+Z=(F+L^{*})T^{-}(F^{*}+L)+[&(I-EE^{-})^{*}V(I-EE^{-})\\ &+(I-K^{-}K)W(I-K^{-}K)^{*}+C-FT^{-}L+C^{*}-L^{*}T^{-}F^{*}].\end{aligned}$$

Observe that $E^{*}(C-FT^{-}L)=L-TT^{-}L=0$ since $\mathcal{M}(L)=\mathcal{M}(T)$ in view of the rank conditions (2.15). Similarly,

$$(C-FT^{-}L)K^{*}=F-FT^{-}T=0.$$

Hence

$$\begin{aligned}C+C^{*}+Y+Z&=(F+L^{*})T^{-}(F^{*}+L)\\ &+[(I-EE^{-})^{*}:(I-K^{-}K)]\begin{pmatrix}V & C-FT^{-}L\\ C^{*}-L^{*}T^{-}F^{*} & W\end{pmatrix}\begin{bmatrix}I-EE^{-}\\ (I-K^{-}K)^{*}\end{bmatrix},\end{aligned}$$

which is nonnegative definite if V and W are so chosen that

$$\begin{pmatrix}V & C-FT^{-}L\\ C^{*}-L^{*}T^{-}F^{*} & W\end{pmatrix}$$

is also. One such choice is $V=I$, $W=(C^{*}-L^{*}T^{-}F^{*})(C-FT^{-}L)$, leading to the following expressions for Y and Z:

$$\begin{aligned}Y&=FT^{-}F^{*}+(I-EE^{-})^{*}(I-EE^{-}),\\ Z&=L^{*}T^{-}L+(C^{*}-L^{*}T^{-}F^{*})(C-FT^{-}L).\end{aligned}$$

With such a choice of Y and Z it is seen that X as in (2.16) satisfies (2.11) and is nonnegative definite.

Let $\hat{X}$ be a nonnegative definite solution of (2.11). Put $Y = A\hat{X}A$, $Z = B\hat{X}B$ and observe that Y, Z are nonnegative definite solutions of (2.14) such that

$$N = C + C^* + Y + Z = (A+B)\hat{X}(A+B)$$

is also nonnegative definite. That $\hat{X}$ can always be expressed as in (2.16) follows from Lemma 2.1.

LEMMA 2.1. *Let A be a given matrix in $\mathscr{C}^{m\times n}$ and C a given nonnegative definite matrix in $\mathscr{C}^{m\times m}$ such that the equation*

$$AXA^* = C \tag{2.17}$$

is consistent. A general nonnegative definite solution to this equation is given by

$$X = A^-C(A^-)^* + (I - A^-A)U(I - A^-A)^*, \tag{2.18}$$

where A^- is an arbitrary g-inverse of A and U is an arbitrary nonnegative definite matrix in $\mathscr{C}^{n\times n}$.

Proof. X as defined in (2.18) is clearly a nonnegative definite solution to (2.17). To show that (2.18) is the expression for a general nonnegative definite solution, we consider an arbitrary such solution $\hat{X}$ and write $\hat{X} = (\hat{Y})(\hat{Y})^*$. Clearly $(A\hat{Y})(A\hat{Y})^* = C$ and therefore rank $A\hat{Y} =$ rank $C = r$ (say). Let us now write $A\hat{Y} = D$ and consider its partitioned form $D = (D_1 \vdots D_2)$, where D_1 is the matrix formed by the first r columns of D. We may now assume, without any loss of generality, that $D_2 = 0$, because if it is not so, one can always realize this objective through a unitary transformation applied to the columns of $\hat{Y}$. If the matrix $\hat{Y}$ is partitioned as $(\hat{Y}_1 \vdots \hat{Y}_2)$ in a corresponding manner, it then follows that

$$A\hat{Y}_1 = D_1, \qquad D_1D_1^* = DD^* = C,$$
$$A\hat{Y}_2 = 0.$$

Since rank $\hat{Y}_1 =$ rank $D_1 = r$, it follows from Note 1 following Lemma 2.2 in Mitra [5] that $\hat{Y}_1 = A^-D_1$ for some g-inverse A^- of A and $\hat{Y}_2 = (I - A^-A)V$ for some V. Hence

$$\hat{X} = \hat{Y}_1\hat{Y}_1^* + \hat{Y}_2\hat{Y}_2^* = A^-C(A^-)^* + (I - A^-A)U(I - A^-A)^*,$$

where $U = VV^*$ is a nonnegative definite matrix. This concludes the proof of Lemma 2.1 and also of Theorem 2.5.

REFERENCES

[1] A. BEN-ISRAEL AND A. BERMAN, *More on linear inequalities with applications to matrix theory*, J. Math. Anal. Appl., 33 (1971), pp. 482–496.

[2] A. BEN-ISRAEL AND T. N. E. GREVILLE, *Generalized Inverses: Theory and Applications*, John Wiley, New York, 1974.

[3] A. BJERHAMMER, *Rectangular reciprocal matrices with special reference to geodetic calculations*, Kungl. Tekn. Hogsk. Handl. Stockholm, 45 (1951), pp. 1–86.

[4] R. D. HILL, *Inertia theorems for simultaneously triangulable complex matrices*, Linear Algebra and Appl., 2 (1969), pp. 131–142.

[5] S. K. MITRA, *Fixed rank solutions of linear matrix equations*, Sankhyā Ser. A, 35 (1972), pp. 387–392.

[6] ——, *Common solutions to a pair of linear matrix equations* $A_1XB_1 = C_1$, $A_2XB_2 = C_2$, Proc. Cambridge Philos. Soc., 74 (1973), pp. 213–216.

[7] G. L. MORRIS AND P. L. ODELL, *Common solutions for n matrix equations with applications*, J. Assoc. Comput. Mach., 15 (1968), pp. 272–274.

[8] R. PENROSE, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

[9] ——, *On best approximate solutions of linear matrix equations*, Ibid., 52 (1956), pp. 17–19.

[10] C. R. RAO, *Generalized inverse for matrices and its applications in mathematical statistics*, Research Papers in Statistics, Festschrift for J. Neyman, John Wiley, New York, 1965.

[11] ——, *Estimation of variance and covariance components in linear models*, J. Amer. Statist. Assoc., 67 (1972), pp. 112–115.

[12] C. R. RAO AND S. K. MITRA, *Generalized Inverse of Matrices and its Applications*, John Wiley, New York, 1971.

[13] G. W. STAGG AND A. H. EL-ABIAD, *Computer Methods in Power System Analysis*, McGraw-Hill, New York, 1968.