

The Matrix Equations $AX = C$, $XB = D$

Sujit Kumar Mitra
Indian Statistical Institute
New Delhi, India

Submitted by Richard A. Brualdi

ABSTRACT

For the pair of matrix equations $AX = C$, $XB = D$ this paper gives common solutions of minimum possible rank and also other feasible specified ranks.

1. INTRODUCTION

In some applications requiring solution of matrix equations one has to seek solution matrices of prescribed ranks. The reflexive generalized inverse (g -inverse) of a matrix A is a matrix X of minimum rank satisfying the equation

$$AXA = A$$

(see e.g., [6, Lemma 2.5.1]). Some g -inverses of A of maximum rank which lead to basic solutions of consistent equations $Ax = y$ have been found useful in linear programming computations ([7]; see also Section 2.8 in [6]). Seshu and Reed [8, Theorem 4-23] show that two nonoriented graphs \mathcal{G}_1 and \mathcal{G}_2 with the respective incidence matrices A_1 and A_2 are 2-isomorphic iff the matrix equation

$$A_2 = XA_1$$

admits a nonsingular solution X . Similar conditions are also involved in the verification whether each row of a matrix F corresponds to a cut set or element disjoint union of cut sets of a graph \mathcal{G} [8, Theorem 4.16].

Keeping such possible applications in view, the author, in an earlier paper [3], obtained solutions of prescribed ranks for the following systems of matrix

equations

$$AX = C, \quad (I)$$

$$AXB = C, \quad (II)$$

where A , B , and C are given matrices. In the present paper, for complex matrices A , B , C , D , and X of appropriate order, we consider the system

$$AX = C, \quad XB = D. \quad (III)$$

Necessary and sufficient conditions for the equations $AX = C$, $XB = D$ to have a common solution were given by Cecioni [2], and the expression for a general common solution by Rao and Mitra [6, p. 25]. For the pair of equations in (III) we obtain a common solution with the minimum possible rank and in fact for any feasible specified rank. The method illustrates another beautiful application of the minimum seminorm g -inverses similar to that in the representation of shorted operators [5].

Interesting byproducts are solutions for the systems (I) and (II) with prescribed ranks for the expression EXF , where E and F are given matrices. It is conceivable that in some applications while solving a matrix equation one may be interested in a particular minor of the solution matrix. One may accordingly stipulate that this minor be of a specified rank. We note that a minor of X can always be expressed in the form EXF for suitable choices of matrices, E and F .

2. RESULTS

Let \mathcal{C}^n and $\mathcal{C}^{m \times n}$ denote respectively the vector spaces of complex n -tuples and complex matrices of order $m \times n$. Let \mathcal{H}_n denote the cone of hermitian nonnegative definite (n.n.d.) matrices of order $n \times n$. For a matrix A , A' denotes its transpose, A^* its complex conjugate transpose, $\mathcal{M}(A)$ its column span, and $\mathcal{N}(A)$ its null space. A^- denotes a generalized inverse (g -inverse) of A , and $A_{m(N)}^-$ a minimum N seminorm g -inverse [6]. The class of minimum N seminorm g -inverses of A is denoted by $\{A_{m(N)}^-\}$. $(A : B)$ denotes a partitioned matrix, the partitioning being understood columnwise. For matrices $A, B \in \mathcal{H}_n$, we write $A > B$ if $A - B \in \mathcal{H}_n$. For $N \in \mathcal{H}_n$ and subspace \mathcal{S} of \mathcal{C}^n , the shorted matrix $\mathcal{S}(N)$ is the unique matrix in \mathcal{H}_n which is such that

$$\begin{aligned} \mathcal{M}(\mathcal{S}(N)) &\subset \mathcal{S}, \\ N &> \mathcal{S}(N), \end{aligned}$$

and if $C \in \mathcal{C}_n$, $\mathcal{M}(C) \subset \mathcal{S}$, and $N > C$, then $\mathcal{S}(N) > C$. The existence of $\mathcal{S}(N)$ was established by Anderson and Trapp [1].

Further, two subspaces of a vector space are said to be virtually disjoint if their intersection consists exclusively of the null vector. We need the following properties of $A_{m(N)}^-$.

THEOREM 1. $G \in \{A_{m(N)}^-\}$ iff

$$AGA = A, \quad (NGA)^* = NGA. \quad (1)$$

THEOREM 2. If $G \in \{A_{m(N)}^-\}$,

- (a) NGA is unique with respect to choice of G in this class.
- (b) $NGA \in \mathcal{C}_n$, $N - NGA \in \mathcal{C}_n$.
- (c) $\mathcal{M}(N - NGA)$ is virtually disjoint with $\mathcal{M}(A^*)$.
- (d) $\mathcal{M}(N) = \mathcal{M}(NGA) \oplus \mathcal{M}(N - NGA)$.
- (e) $\mathcal{M}(NGA) = \mathcal{M}(N) \cap \mathcal{M}(A^*)$.
- (f) If $\mathcal{S} = \mathcal{M}(A^*)$ then $NGA = \mathcal{S}(N)$.

Theorem 1 is proved in Rao and Mitra [6, p. 46], and Theorem 2 in Mitra and Puri [5].

We next prove a few lemmas which are also needed. The lemmas are also of independent interest. Let $A, E \in \mathcal{C}_m$, $B, F \in \mathcal{C}_n$ and the equation

$$AXB = C \quad (2)$$

be consistent. Let Ω denote the class of solutions of (2).

LEMMA 1. $\min_{X \in \Omega} \text{rank } EXF = \text{rank } EA_{m(E)}^- C [B_{m(F)}^-]^* F$, and the minimum is attained by $X = A_{m(E)}^- C [B_{m(F)}^-]^* \in \Omega$.

Proof. For $X \in \Omega$

$$\begin{aligned} EA_{m(E)}^- C [B_{m(F)}^-]^* F &= EA_{m(E)}^- AXB [B_{m(F)}^-]^* F \\ &= A(A_{m(E)}^-)^* EXFB_{m(F)}^- B, \end{aligned}$$

on account of (1). From this Lemma 1 follows, since $A_{m(E)}^- C [B_{m(F)}^-]^*$ clearly is a solution of (2).

Let $A \in \mathcal{C}^{p \times m}$, $B \in \mathcal{C}^{n \times q}$, $C \in \mathcal{C}^{p \times q}$, $E \in \mathcal{C}^{s \times m}$, $F \in \mathcal{C}^{n \times t}$, and the equation $AXB = C$ be consistent. Let Ω denote the class of solutions. As a

simple corollary to Lemma 1 we have Lemma 2. See also Mitra [4] in this connection.

LEMMA 2. $\min_{X \in \Omega} \text{rank } EXF = \text{rank } E^*E(A^*A)^{-}_{m(E^*E)}A^*CB^*$
 $((BB^*)^{-}_{m(F^*F)})^*FF^*$, and this minimum is attained by

$$X = (A^*A)^{-}_{m(E^*E)}A^*CB^*[(BB^*)^{-}_{m(F^*F)}]^* \in \Omega. \quad (3)$$

LEMMA 3. Let $N \in \mathcal{C}_n$, $\mathcal{S} \subset \mathcal{E}^n$, and $\mathcal{T} \subset \mathcal{M}[N - \mathcal{S}(N)]$. Then

$$(\mathcal{S} \oplus \mathcal{T})(N) = \mathcal{S}(N) + \mathcal{T}(N), \quad (4)$$

where $\mathcal{S}(N)$, $\mathcal{T}(N)$, and $(\mathcal{S} \oplus \mathcal{T})(N)$ denote the shorted versions of N with reference to subspaces \mathcal{S} , \mathcal{T} , and $\mathcal{S} \oplus \mathcal{T}$ respectively.

Proof. Write $N_0 = (\mathcal{S} \oplus \mathcal{T})(N)$, and observe that by definition of a shorted operator $N > N_0$. Also $\mathcal{S}(N_0) = \mathcal{S}[(\mathcal{S} \oplus \mathcal{T})(N)] = \{\mathcal{S} \cap (\mathcal{S} \oplus \mathcal{T})\}(N) = \mathcal{S}(N)$ by Corollary 5 to Theorem 1 in [1]. Hence

$$\begin{aligned} N - \mathcal{S}(N) &> N_0 - \mathcal{S}(N_0) \\ &\Rightarrow \mathcal{M}[N_0 - \mathcal{S}(N_0)] \subset \mathcal{M}[N - \mathcal{S}(N)]. \end{aligned}$$

We next observe that, on account of (1) and Theorem 2(f),

$$y^*N^-x = 0 \quad \forall x \in \mathcal{M}[N - \mathcal{S}(N)], y \in \mathcal{M}[\mathcal{S}(N)] \text{ and } \forall N^-.$$

Hence

$$y^*N^-x = 0 \quad \forall x \in \mathcal{M}[N_0 - \mathcal{S}(N_0)], y \in \mathcal{M}[\mathcal{S}(N)] \text{ and } \forall N^-.$$

Further, $\mathcal{M}(N_0) = \{\mathcal{S} \oplus \mathcal{T}\} \cap \mathcal{M}(N) = \{\mathcal{S} \cap \mathcal{M}(N)\} \oplus \mathcal{T}$ by Theorem 2(e), since $\mathcal{T} \subset \mathcal{M}(N)$. Hence $\mathcal{M}[N_0 - \mathcal{S}(N_0)] \subset \mathcal{M}[\mathcal{S}(N)] \oplus \mathcal{T}$. Let x be an arbitrary vector in $\mathcal{M}[N_0 - \mathcal{S}(N_0)]$. Write $x = x_1 + x_2$, where $x_1 \in \mathcal{T}$ and $x_2 \in \mathcal{M}[\mathcal{S}(N)]$. Then $y^*N^-x = y^*N^-x_2 = 0 \quad \forall y \in \mathcal{M}[\mathcal{S}(N)] \Rightarrow \mathcal{S}(N) \perp x_2 = x_2 = 0$, since $N^- \in \{(\mathcal{S}(N))^\perp\}$ (by Theorem 2.4 in [5]), and $x_1 \in \mathcal{M}[\mathcal{S}(N)] \Rightarrow x = x_1 \Rightarrow \mathcal{M}[N_0 - \mathcal{S}(N_0)] \subset \mathcal{T} \Rightarrow \mathcal{T}[N_0 - \mathcal{S}(N_0)] = N_0^- - \mathcal{S}(N_0) < \mathcal{T}(N_0) < \mathcal{T}(N)$. Unless equality holds here, $\exists K \in \mathcal{C}_n$ such that

$K \neq 0$, $\mathcal{M}(K) \subset \mathcal{F}$, and

$$N_0 = \mathcal{S}(N) + \mathcal{T}(N) - K.$$

Since $\mathcal{M}(K) \not\subset \mathcal{M}[\mathcal{S}(N)]$, $\exists u \in \mathcal{C}^n$ such that $u \in \mathcal{N}[\mathcal{S}(N)]$, $u \notin \mathcal{N}(K)$, and

$$u^* N_0 u = u^* \mathcal{S}(N) u - u^* K u,$$

which is strictly less than $u^* \mathcal{F}(N) u$. This contradicts the inequality

$$N_0 > \mathcal{F}(N_0) = \mathcal{F}[\{\mathcal{S} \oplus \mathcal{T}\}(N)] = [\{\mathcal{S} \oplus \mathcal{T}\} \cap \mathcal{F}](N) = \mathcal{F}(N).$$

Hence $N_0 - \mathcal{S}(N_0) = \mathcal{T}(N)$ and Lemma 3 is established.

That (4) is not true in general can be seen for $n = 2$ with

$$N = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathcal{S} = \mathcal{M} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathcal{F} = \mathcal{M} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Note that here (4) is not true even though \mathcal{S} and \mathcal{F} are virtually disjoint.

Let $A \in \mathcal{C}^{p \times m}$, $B \in \mathcal{C}^{n \times q}$, $C \in \mathcal{C}^{p \times n}$, $D \in \mathcal{C}^{m \times q}$, and the equations $AX = C$ and $XB = D$ be individually consistent. Assume further

$$AD = CB, \tag{5}$$

which is both necessary and sufficient for the pair of equations

$$AX = C, \quad XB = D \tag{6}$$

to have a common solution. Assume without loss of generality that

$$\text{rank } C \leq \text{rank } D. \tag{7}$$

If X is a common solution,

$$\text{rank } X \geq \max\{\text{rank } C, \text{rank } D\} = \text{rank } D. \tag{8}$$

THEOREM 3. *The pair of equations (6) have a common solution of rank equal to rank D iff*

$$\text{rank } CB = \text{rank } C. \tag{9}$$

Proof. If X is a common solution, by the Frobenius inequality

$$\begin{aligned} \text{rank } AXB + \text{rank } X &\geq \text{rank } AX + \text{rank } XB \\ \Rightarrow \text{rank } X &\geq \text{rank } C + \text{rank } D - \text{rank } CB, \end{aligned}$$

From this, the necessity of the condition (9) follows.

If (9) holds, $Y = C$ is a solution of the equation

$$YB = AD$$

of rank equal to rank AD . Hence by Note 1 following Lemma 2.2 in Mitra [3], there exists a g -inverse $(B^-)_0$ of B such that

$$C = AD(B^-)_0.$$

Clearly $X = D(B^-)_0$ is a common solution of rank equal to rank D . This concludes proof of sufficiency part and of Theorem 3.

We next consider the case where

$$\text{rank } C - \text{rank } CB = \delta > 0.$$

Here let CB_0 be a matrix of δ linearly independent columns such that

$$\mathcal{M}(C) = \mathcal{M}(CB) \oplus \mathcal{M}(CB_0). \quad (10)$$

Since $\mathcal{M}(C) \subset \mathcal{M}(A)$, the equation

$$AY = CB_0 \quad (11)$$

is consistent. Let $Y = K$ be a solution. Clearly the δ columns of B_0 are linearly independent. This implies that the equation $XB_0 = K$ is consistent, which together with the consistency of $XB = D$ and on account of (10) implies the consistency of

$$X(B; B_0) = (D; K).$$

Also

$$A(D; K) = C(B; B_0)$$

and

$$\text{rank } C(B; B_0) = \text{rank } C.$$

This implies that the pair of equations

$$AX = C, \quad X(B; B_0) = (D; K) \quad (12)$$

satisfy the conditions of Theorem 3 and therefore have a common solution of minimum possible rank equal to $\text{rank}(D; K)$. Noting that each common solution to the pair of equations (6) corresponds to a matrix K which also satisfies (11), we see that the problem of finding a minimum rank common solution to the pair of equations (6) reduces to that of finding a solution Y of (11) such that $\text{rank}(D; Y)$ is minimum, or equivalently $\text{rank } EY$ is minimum, where

$$\mathcal{M}(E^*) = \mathcal{N}(D^*),$$

since for such a choice of E , $\text{rank } EY = \text{rank}(D; Y) - \text{rank } D$. (See e.g. Lemma 7.1.2 of Rao and Mitra [6], which is precisely the same result for the real case. The proof for the complex case is similar.)

Assume now that E is n.n.d. One such choice of E is given by

$$E = I - D(D^*D)^- D^*.$$

By Lemma 2, the required choice for Y is given by

$$Y = (A^*A)_{m(E)}^- A^*CB_0 = K_0 \quad (\text{say}), \quad (13)$$

and the minimum possible rank for a common solution to the pair of equations (6) is

$$\text{rank}(D; (A^*A)_{m(E)}^- A^*CB_0).$$

We have thus arrived at the following theorem.

THEOREM 4. *Let $\text{rank } CB < \text{rank } C \leq \text{rank } D$. Let B_0 be determined as in (10). If X is a common solution to the pair of equations (6),*

$$\text{rank } X \geq \text{rank}(D; (A^*A)_{m(E)}^- A^*CB_0). \quad (14)$$

Further, the lower bound is attainable.

A common solution of minimum possible rank in this case is obtained as indicated in proof of Theorem 3, as applied to (12) with K replaced by K_0 defined in (13).

Assume now that (9) holds, so that we are in the situation covered by Theorem 3. In terms of a particular common solution $D(B^-)_0$ to the equations in (6), a general common solution is given by

$$X = D(B^-)_0 + (I - A^-A)Z(I - B(B^-)_0), \quad (15)$$

where $Z \in \mathcal{C}^{m \times n}$ and is arbitrary (Theorem 2.33 of [6]). For reasons which will be clear shortly, we choose $A_{m(t)}^-$ for A^- so that

$$I - A^-A = I - A_{m(t)}^-, A = I - A^*(AA^*)^-A = Q$$

is n.n.d. Let us rewrite (15) as

$$\begin{aligned} X &= D(B^-)_0 + Q(D^*)_{m(Q)}^- D^* Z(I - B(B^-)_0) \\ &\quad + Q(I - (D^*)_{m(Q)}^- D^*) Z(I - B(B^-)_0) \\ &= D(B^-)_1 + Q(I - (D^*)_{m(Q)}^- D^*) Z(I - B(B^-)_0), \end{aligned}$$

where $(B^-)_1 = (B^-)_0 + [(D^*)_{m(Q)}^-]^* Q Z(I - B(B^-)_0) \in \{B^-\}$.

Since the row spans of $D(B^-)_1$ and $I - B(B^-)_0$ are virtually disjoint and by Theorem 2(c) $\mathcal{M}\{Q(I - (D^*)_{m(Q)}^- D^*)\}$ is virtually disjoint from $\mathcal{M}(D)$, we have

rank X

$$\begin{aligned} &= \text{rank } D(B^-)_1 + \text{rank} \{Q(I - (D^*)_{m(Q)}^- D^*) Z(I - B(B^-)_0)\} \\ &\leq \text{rank } D(B^-)_1 + \min \{ \text{rank } Q(I - (D^*)_{m(Q)}^- D^*), \text{rank} [I - B(B^-)_0] \} \\ &= \text{rank } D + \min \{ \text{rank}(Q : D) - \text{rank } D, n - \text{rank } B \} \text{ using Theorem 2(d)} \\ &= \min \{ \text{rank}(Q : D), n - \text{rank } B + \text{rank } D \} \\ &= \min \{ m - \text{rank } A + \text{rank } AD, n - \text{rank } B + \text{rank } D \} \\ &= \min \{ m - \text{rank } A + \text{rank } C, n - \text{rank } B + \text{rank } D \} = \theta \quad (\text{say}), \quad (16) \end{aligned}$$

since

$$\begin{aligned} \text{rank}(Q: D) &= \text{rank}(Q: DD^*) = \text{rank}(Q: DD^*A^*) \\ &\quad + \text{rank } Q + \text{rank } DD^*A^* \\ &= m - \text{rank } A + \text{rank } AD. \end{aligned} \quad (17)$$

For rank $D < s < \theta$, X will be a common solution of rank equal to s iff Z is so chosen that

$$\text{rank } Q(I - (D^*)_{m(Q)}^-)Z(I - B(B^-)_0) = s - \text{rank } D. \quad (18)$$

A general common solution of rank s is given by

$$\begin{aligned} X &= D(B^-)_0 + Q(D^*)_{m(Q)}^- D^* Z_0 (I - B(B^-)_0) \\ &\quad + Q(I - (D^*)_{m(Q)}^-)Z(I - B(B^-)_0), \end{aligned} \quad (19)$$

where $Z_0 \in \mathcal{C}^{m \times n}$ is arbitrary and $Z \in \mathcal{C}^{m \times n}$ satisfies the condition (18) but is otherwise arbitrary. A general common solution of rank s in the general case when (9) is not true can be obtained in a like manner with the help of Theorem 4.

For completeness we describe here a method of obtaining a solution X of the consistent equation $AXB = C$ such that the matrix EXF has a specified rank. We confine our attention to the case where the coefficient matrices A , B , E , and F are hermitian n.n.d. as considered in Lemma 1. The general case corresponding to Lemma 2 can be treated in a like manner. Let δ_1 be an integer,

$$\delta_1 \leq \min\{\text{rank } E(I - A_{m(E)}^-), \text{rank } F(I - B_{m(F)}^-)B\},$$

and A_0, B_0 be matrices of rank δ_1 in $\mathcal{C}^{m \times \delta_1}$ and $\mathcal{C}^{n \times \delta_1}$, respectively such that $\mathcal{M}(A_0) \subset \mathcal{M}\{E(I - A_{m(E)}^-)A\}$, $\mathcal{M}(B_0) \subset \mathcal{M}\{F(I - B_{m(F)}^-)B\}$. Put $A_1 = A_0 A_0^*$, $B_1 = B_0 B_0^*$, $C_1 = A_0 B_0^*$, and consider the equation

$$(A + A_1)X(B + B_1) = C + C_1. \quad (20)$$

This equation is clearly consistent, and any solution of it is a solution of (3). A solution of (20) which gives the minimum possible rank for EXF is, by Lemma

1, given by

$$X_0 = (A + A_1)_{m(E)}^-(C + C_1)[(B + B_1)_{m(F)}^-]^*.$$

Further, by Lemma 3,

$$\begin{aligned} \text{rank}(EX_0F) &= \text{rank}\left\{E(A + A_1)_{m(E)}^-C[(B + B_1)_{m(F)}^-]^*F\right. \\ &\quad \left.+ E(A + A_1)_{m(E)}^-C_1[(B + B_1)_{m(F)}^-]^*F\right\} \\ &= \text{rank}\left\{EA_{m(E)}^-C[(B)_{m(F)}^-]^*F + E(A_1)_{m(E)}^-C_1[(B_1)_{m(F)}^-]^*F\right\} \\ &= \delta + \delta_1, \end{aligned}$$

where $\delta = \text{rank}\{EA_{m(E)}^-C[(B)_{m(F)}^-]^*F\}$, since if $\mathcal{S} = \mathcal{A}(A)$, $\mathcal{T} = \mathcal{A}(A_1)$, then

$$\begin{aligned} (\mathcal{S} \oplus \mathcal{T})(E) &= E(A + A_1)_{m(E)}^-(A + A_1) \\ &= E(A + A_1)_{m(E)}^-A + E(A + A_1)_{m(E)}^-A_1 \\ &= \mathcal{S}(E) + \mathcal{T}(E) = E(A)_{m(E)}^-A + E(A_1)_{m(E)}^-A_1 \\ \Rightarrow E(A + A_1)_{m(E)}^-A &= EA_{m(E)}^-A, \quad E(A + A_1)_{m(E)}^-A_1 = E(A_1)_{m(E)}^-A_1. \end{aligned}$$

Similarly $F(B + B_1)_{m(F)}^-B = FB_{m(F)}^-B$, $F(B + B_1)_{m(F)}^-B_1 = F(B_1)_{m(F)}^-B_1$.

Thanks are due to Thomas Mathew for providing the simplification given in (17).

REFERENCES

- 1 W. N. Anderson, Jr., and G. Trapp, Shorted operators II, *SIAM J. Appl. Math.* 28:60-71 (1975).
- 2 F. Cecioni, Sopra operazioni algebriche, *Ann. Scuola Norm. Sup. Pisa Sci. Fis. Mat.* 11:17-20 (1910).
- 3 S. K. Mitra, Fixed rank solutions of linear matrix equations, *Sankhyā Ser. A* 35:387-392 (1972).

- 4 S. K. Mitra, Common solutions to a pair of linear matrix equations $A_1XB_1 = C_1$, $A_2XB_2 = C_2$, *Proc. Cambridge Philos. Soc.* 74:213-216 (1973).
- 5 S. K. Mitra and M. L. Puri, Shorted operators and generalized inverses of matrices, *Linear Algebra Appl.* 25:45-56 (1979).
- 6 C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.
- 7 J. B. Rosen, Minimum and basic solutions to singular linear systems, *SIAM J. Appl. Math.* 12:156-162 (1964).
- 8 S. Seshu and M. B. Reed, *Linear Graphs and Electrical Networks*, Addison-Wesley, Reading, Mass., 1961.

Received 30 December 1982; revised 25 April 1983