Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

Jerzy K. Baksalary

The pair of matrix equations $AX = B$ and $A^*Y + CX = D$

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **73** (1982), n.5, p. 81–88.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1982_8_73_5_81_0>

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Accademia Nazionale dei Lincei, 1982.

RENDICONTI

DELLE SEDUTE

DELLA ACCADEMIA NAZIONALE DEI LINCEI

Classe di Scienze fisiche, matematiche e naturali

Seduta del 25 novembre 1982

Presiede il Presidente della Classe GIUSEPPE MONTALENTI

SEZIONE I

(Matematica, meccanica, astronomia, geodesia e geofisica)

Algebra. — *The pair of matrix equations* $AX = B$ *and* $A^*Y + CX = D$ (*). Nota di JERZY K. BAKSALARY (**), presentata (***) dal Corrisp. M. AGENO.

RIASSUNTO. — È stata trovata una condizione necessaria e sufficiente affinchè le due equazioni matriciali $AX = B$ e $A^*Y + CX = D$ ammettano una soluzione comune. Di quest'ultima è poi stata data una rappresentazione generale, per il caso in cui la condizione trovata sia soddisfatta. È stato inoltre formulato un criterio per l'unicità della soluzione e, ove essa sia unica, ne è stata determinata la forma.

Questi problemi erano stati in precedenza trattati da V. Valerio (1976), ma le conclusioni a cui egli è pervenuto non sono corrette.

1. INTRODUCTION AND PRELIMINARIES

Let $\mathbf{C}_{m,n}$ denote the vector space of $m \times n$ matrices over the complex field. Given $A \in \mathbf{C}_{m,n}$, the symbols A^*, $\mathbf{R}(A)$, $\mathbf{N}(A)$, and $r(A)$ will stand for the conjugate transpose, range, null space, and rank, respectively, of A. Furthermore, A^- will denote a generalized inverse of A, that is, any solution to the equation

$$AA^-A = A,$$

(*) This work was partially supported by the Polish Academy of Sciences Grant No. MR I.1 – 2/1.

(**) Department of Mathematical and Statistical Methods, Academy of Agriculture, 60–637 Poznan, Poland.

(***) Nella seduta del 25 novembre 1982.

while A^+ will denote the Moore–Penrose inverse of A, that is, the unique solution to the set of equations

$$AA^+A = A \quad , \quad A^+AA^+ = A^+ \quad , \quad (AA^+)^* = AA^+, \qquad \text{and} \quad (A^+A)^* = A^+A.$$

It is known that $P_A = AA^+$ is the projector onto $\mathbf{R}(A)$ along $\mathbf{N}(A^*)$, and that $Q_A = I_m - P_A$ (I_m stands for the identity matrix of order m) is the projector onto $\mathbf{N}(A^*)$ along $\mathbf{R}(A)$. Obviously, if $A \in \mathbf{C}_{m,m}$ and is nonsingular, then $A^+ = A^{-1}$, thus implying that $P_A = I_m$ and $Q_A = O_{m,m}$, where $O_{m,m}$ denotes the null matrix of size $m \times m$.

The present note deals with the pair of linear matrix equations,

$$AX = B, \tag{1}$$

$$A^*Y + CX = D, \tag{2}$$

wherein $A \in \mathbf{C}_{m,n}$, $B \in \mathbf{C}_{m,p}$, $C \in \mathbf{C}_{n,n}$, and $D \in \mathbf{C}_{n,p}$ are known. A criterion for the consistency of this pair is derived, and, if a common solution exists, its general representation is given. Moreover, a criterion for the uniqueness of the common solution is established, and an explicit form of the unique solution is found. The pair of equations (1) and (2) has already been examined by Valerio [8], especially in the context of some problems concerning reticulated structures, but the statements given by him as the main results are false, which will be exhibited by providing appropriate counterexamples.

The developments of this note are essentially based on the following well known results pertaining to a simple matrix equation of the form

$$MX = N, \tag{3}$$

with known $M \in \mathbf{C}_{s,t}$ and $N \in \mathbf{C}_{s,u}$.

LEMMA 1. *Equation* (3) *is consistent if and only if*

$$\mathbf{R}(N) \subset \mathbf{R}(M).$$

LEMMA 2. *If equation* (3) *is consistent, then its general solution is expressible as*

$$X = M^- N + (I_t - M^- M) Z,$$

where Z *varies over* $\mathbf{C}_{t,u}$.

LEMMA 3. *If equation* (3) *is consistent, then it admits the unique solution* $X = M^- N$ *if and only if* $r(M) = t$, *in which case the product* $M^- N$ *is independent of the choice of* M^-.

2. Consistency and a general representation of the solution

In his Theorem 1, Valerio [8] states that the pair of equations (1) and (2) is consistent if and only if the equation (1) is consistent, which, on account of Lemma 1, is equivalent to the inclusion $\mathbf{R}(B) \subset \mathbf{R}(A)$. This statement is rather curious, for even the consistency of each of the equations (1) and (2) is not, in general, a sufficient condition for the consistency of the pair comprising (1) and (2). Taking, for example,

$$\mathrm{A} = \mathrm{D}^* = (1 \quad 0) \quad , \quad \mathrm{B} = (1) , \qquad \text{and} \quad \mathrm{C} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

it is easy to verify that no $\mathrm{X} \in \mathbf{C}_{1,2}$ and $\mathrm{Y} \in \mathbf{C}_{1,1}$ exist which would satisfy (1) and (2) simultaneously, in spite of the fact that the conditions $\mathbf{R}(\mathrm{B}) \subset \mathbf{R}(\mathrm{A})$ and $\mathbf{R}(\mathrm{D}) \subset \mathbf{R}(\mathrm{C} \,\vdots\, \mathrm{A}^*)$ are fulfilled. A correct consistency criterion is revealed in the following

THEOREM 1. *The pair of equations* (1) *and* (2) *admits a common solution if and only if*

$$\mathbf{R}(\mathrm{B}) \subset \mathbf{R}(\mathrm{A}) \tag{4}$$

and

$$\mathbf{R}(\mathrm{D} - \mathrm{CA}^+ \mathrm{B}) \subset \mathbf{R}(\mathrm{A}^* \,\vdots\, \mathrm{CQ}_{\mathrm{A}^*}) . \tag{5}$$

Proof. On account of Lemma 1, the equation (1) is consistent if and only if (4) holds, in which case, according to Lemma 2, the general solution may be written as

$$\mathrm{X} = \mathrm{A}^+ \mathrm{B} + \mathrm{Q}_{\mathrm{A}^*} \mathrm{V} , \tag{6}$$

with V varying over $\mathbf{C}_{n,p}$. Substituting (6) into equation (2) transforms the latter to the form

$$(\mathrm{A}^* \,\vdots\, \mathrm{CQ}_{\mathrm{A}^*}) \begin{pmatrix} \mathrm{Y} \\ \mathrm{V} \end{pmatrix} = \mathrm{D} - \mathrm{CA}^+ \mathrm{B} . \tag{7}$$

Hence it follows that condition (4) is to be supplemented by a necessary and sufficient condition for the consistency of (7) which, in view of Lemma 1, expresses as in (5). The proof is complete.

Utilizing remarks given on p. 517 of [1], it can be noted that an equivalent form of (5) is

$$\mathbf{R}(\mathrm{D} - \mathrm{CA}^+ \mathrm{B}) \subset \mathbf{R}(\mathrm{A}^*) \boxed{+} \mathbf{R}(\mathrm{Q}_{\mathrm{A}^*} \mathrm{CQ}_{\mathrm{A}^*}) ,$$

where $\boxed{+}$ denotes the orthogonal sum of the specified subspaces. On the other hand, Rao's [6] Lemma 2.1 assures that if C is nonnegative definite, then

$$\mathbf{R}(\mathrm{A}^* \,\vdots\, \mathrm{CQ}_{\mathrm{A}^*}) = \mathbf{R}(\mathrm{A}^* \,\vdots\, \mathrm{C}) , \tag{8}$$

thus implying a simplification of (5) to the form

$$\mathbf{R}(D) \subset \mathbf{R}(A^* : C). \tag{9}$$

Observing that (9) is a necessary and sufficient condition for the consistency of (2) leads to the following

COROLLARY 1. *The pair of equations* (1) *and* (2), *with a nonnegative definite* C, *admits a common solution if and only if each of these equations admits a solution.*

Having determined a criterion for the existence of a common solution to the considered pair of equations, it is now natural to ask about its form.

THEOREM 2. *If the pair of equations* (1) *and* (2) *admits a common solution, then a general representation of the solution is*

$$X = A^+ B + S^+ (D - CA^+ B) + Q_{(A^* : S^*)} W \tag{10}$$

and

$$Y = A^{*+} (I_n - CS^+)(D - CA^+ B) - A^{*+} CQ_{(A^* : S^*)} W + Q_A Z , \tag{11}$$

where

$$S = Q_{A^*} CQ_{A^*} , \tag{12}$$

while W *and* Z *vary over* $\mathbf{C}_{n,p}$ *and* $\mathbf{C}_{m,p}$, *respectively.*

Proof. From the proof of Theorem 1 it is clear that the crucial point in developing a general representation of the common solution to equations (1) and (2) is to devise a solution to equation (7). On account of Theorem 3.1 in Pringle and Rayner [5], one of the generalized inverses of $(A^* : CQ_{A^*})$ is

$$(A^* : CQ_{A^*})^- = \begin{pmatrix} A^{*+} - A^{*+} CQ_{A^*} S^+ Q_{A^*} \\ S^+ Q_{A^*} \end{pmatrix}.$$

But according to a result on p. 682 of [3] (see also Lemma 2 in [1]),

$$Q_{A^*} S^+ = S^+ = S^+ Q_{A^*} , \tag{13}$$

and, therefore,

$$(A^* : CQ_{A^*})^- = \begin{pmatrix} A^{*+} (I_n - CS^+) \\ S^+ \end{pmatrix}.$$

Hence, using (13) and the fact that $S^+ A^* = O_{n,m}$,

$$(A^* : CQ_{A^*})^- (A^* : CQ_{A^*}) = \begin{pmatrix} P_A & A^{*+} CQ_{A^*} Q_{S^*} \\ O_{n,m} & P_{S^*} \end{pmatrix}.$$

In view of Lemma 2, it now follows that a general representation of the solution to equation (7) is expressible as

$$Y = A^{*+}(I_n - CS^+)(D - CA^+ B) + Q_A Z - A^{*+} CQ_{A^*} Q_{S^*} W \tag{14}$$

and

$$V = S^+ (D - CA^+ B) + Q_{S^*} W, \tag{15}$$

where W and Z vary over $\mathbf{C}_{n,p}$ and $\mathbf{C}_{m,p}$, respectively. Substituting (15) into (6) and applying (13) gives

$$X = A^+ B + S^+ (D - CA^+ B) + Q_{A^*} Q_{S^*} W. \tag{16}$$

Finally, observe that

$$Q_{A^*} Q_{S^*} = I_n - (P_{A^*} + P_{S^*}). \tag{17}$$

Since

$$P_{A^*} P_{S^*} = P_{S^*} P_{A^*} = O_{n,n},$$

it follows by Theorem 5.1.2 in Rao and Mitra [7] that

$$P_{A^*} + P_{S^*} = P_{(A^* : S^*)}.$$

Hence, in view of (17),

$$Q_{A^*} Q_{S^*} = Q_{(A^* : S^*)},$$

which transforms (14) and (16) to the required forms (10) and (11), and thus completes the proof.

Repeating the arguments following Theorem 1, it can be noted that the projector $Q_{(A^* : S^*)}$, occurring in (10) and (11), is alternatively expressible as $Q_{(A^* : CQ_{A^*})}$, and also that a substantial simplication of a general representation of the common solution is achieved when C is nonnegative definite, in which case, according to (8), $Q_{(A^* : S^*)} = Q_{(A^* : C^*)}$, and, consequently,

$$A^{*+} CQ_{(A^* : S^*)} = O_{m,n},$$

thus leading to the following

COROLLARY 2. *If the pair of equations* (1) *and* (2), *with a nonnegative definite* C, *admits a common solution, then a general representation of the solution is*

$$X = A^+ B + S^+ (D - CA^+ B) + Q_{(A^* : C^*)} W$$

and

$$Y = A^{*+} (I_n - CS^+)(D - CA^+ B) + Q_A Z,$$

where W *and* Z *vary over* $\mathbf{C}_{n,p}$ *and* $\mathbf{C}_{m,p}$, *respectively.*

3. Uniqueness

In the present section, the considered pair of equations is constantly assumed to be consistent. In his Theorem 2, Valerio [8] states that a necessary and sufficient condition for the uniqueness of the common solution to the equations (1) and (2) is that A and C be both of full rank, or, more precisely, that $r(\mathrm{A}) = m$ and $r(\mathrm{C}) = n$. It can be seen, however, that this statement is false. An example wherein

$$\mathrm{A} = (1 \quad 1) \quad , \quad \mathrm{B} = (0) \quad , \quad \mathrm{C} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} , \qquad \text{and } \mathrm{D} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

shows that Valerio's condition is not necessary, while an example wherein

$$\mathrm{A} = (1 \quad 0) \quad , \quad \mathrm{B} = (0) \quad , \quad \mathrm{C} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \text{and } \mathrm{D} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

shows that it is not sufficient. The first of these examples also exhibits the incorrectness of an explicit representation of the unique common solution, provided by Valerio in his formulas (9′) and (9″). These formulas fail in any case where C is singular.

A correct criterion for the uniqueness of the common solution to equations (1) and (2) can be obtained by putting the requirement that the general representations of X and Y, as given in (10) and (11), be independent of the choice of $\mathrm{W} \in \mathbf{C}_{n,p}$ and $\mathrm{Z} \in \mathbf{C}_{m,p}$. It is clear that this holds if and only if

$$\mathrm{Q}_{(\mathrm{A}^* \,:\, \mathrm{S}^*)} = \mathrm{O}_{n,n} \qquad \text{and} \quad \mathrm{Q}_{\mathrm{A}} = \mathrm{O}_{m,m} ,$$

or, in terms of the ranks of matrices, if and only if

$$r(\mathrm{A}^* : \mathrm{S}^*) = n \qquad \text{and} \quad r(\mathrm{A}) = m . \tag{18}$$

Since, on account of Theorem 5 in Marsaglia and Styan [4], $r(\mathrm{A}^* : \mathrm{S}^*) = r(\mathrm{A}) + r(\mathrm{S})$, the first of the equalities in (18) may be replaced by $r(\mathrm{S}) = n - m$. An explicit form of the unique common solution follows immediately from (10) and (11), thus concluding the proof of the following

Theorem 3. *If the pair of equations* (1) *and* (2) *admits a common solution, the solution is unique if and only if*

$$r(\mathrm{A}) = m \tag{19}$$

and

$$r(\mathrm{Q}_{\mathrm{A}^*} \mathrm{C} \mathrm{Q}_{\mathrm{A}^*}) = n - m . \tag{20}$$

If this is the case, then the unique solution is

$$X = A^+ B + S^+ (D - CA^+ B)$$

and

$$Y = A^{*+} (I_n - CS^+) (D - CA^+ B),$$

with S *as defined in* (12).

It is obvious that the condition (19) is equivalent to $\mathbf{N}(A^*) = \{O_{m,1}\}$. On the other hand, since $r(Q_{A^*}) = n - r(A)$, it is clear that the condition (20) may be written as

$$r(Q_{A^*} C Q_{A^*}) = r(Q_{A^*}). \tag{21}$$

From Corollary 6.2 in Marsaglia and Styan [4] it follows that (21) holds if and only if

$$\mathbf{N}(Q_{A^*} C) \cap \mathbf{R}(Q_{A^*}) = \{O_{n,1}\},$$

or, equivalently, if and only if

$$\mathbf{N}(Q_{A^*} C) \cap \mathbf{N}(A) = \{O_{n,1}\}.$$

These observations are summarized below to form a geometrical version of the criterion for the uniqueness of the common solution to equations (1) and (2).

Remark. The rank conditions (19) and (20) involved in Theorem 3 are equivalent to the relations

$$\mathbf{N}(A^*) = \{O_{m,1}\} \quad \text{and} \quad \mathbf{N}(Q_{A^*} C) \cap \mathbf{N}(A) = \{O_{n,1}\},$$

as well as to the relations

$$\mathbf{R}(A) = \mathbf{C}_{m,1} \quad \text{and} \quad \mathbf{R}(A^* : C^* Q_{A^*}) = \mathbf{C}_{n,1}. \tag{22}$$

It may be pointed out that some other alternative formulations of the original condition (20) are available by applying results given on p. 88 in Ben–Israel and Greville [2] to (21). On the other hand, it seems noticeable that the uniqueness criterion simplifies when C is nonnegative definite. In fact, on account of (8), the second condition in (22) reduces to

$$\mathbf{R}(A^* : C) = \mathbf{C}_{n,1},$$

or, in other words, to

$$\mathbf{N}(A) \cap \mathbf{N}(C) = \{O_{n,1}\},$$

thus leading to the following

COROLLARY 3. *If the pair of equations* (1) *and* (2), *with a nonnegative definite* C, *admits a common solution, the solution is unique if, and only if,* A *and*

$(A^* \vdots C)$ *both are of full row rank, or, equivalently, if, and only if, the null space of* A^* *and the intersection of the null spaces of* A *and* C *both contain merely null vectors.*

It should be mentioned, in conclusion, that some alternatives to the results developed in the present note can be obtained by rewriting the pair of equations (1) and (2) as a single linear matrix equation of the form

$$\begin{pmatrix} C & A^* \\ A & O_{m,m} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} D \\ B \end{pmatrix}, \tag{23}$$

to which Lemmas 1, 2 and 3 apply directly. The crucial point in such an approach is to find an explicit representation of a generalized inverse of the partitioned matrix specified in (23). Clearly, any such generalized inverse is the ordinary inverse if and only if the solution to (23) is unique.

References

[1] J. K. Baksalary and R. Kala (1978) – *Relationships between some representations of the best linear unbiased estimator in the general Gauss–Markoff model,* « SIAM J. Appl. Math. », *35,* 515–520.

[2] A. Ben-Israel and T. N. E. Greville (1974) – *Generalized Inverses: Theory and Applications,* John Wiley, New York.

[3] R. E. Cline and T. N. E. Greville (1970) – *An extension of the generalized inverse of a matrix,* « SIAM J. Appl. Math. », *19,* 682–688.

[4] G. Marsaglia and G. P. H. Styan (1974) – *Equalities and inequalities for ranks of matrices,* « Linear and Multilinear Algebra », *2,* 269–292.

[5] R. M. Pringle and A. A. Rayner (1971) – *Generalized Inverse Matrices with Applications to Statistics,* Griffin, London.

[6] C. R. Rao (1974) – *Projectors, generalized inverses and the BLUE's,* « J. Roy. Statist. Soc. Ser. B », *36,* 442–448.

[7] C. R. Rao and S. K. Mitra (1971) – *Generalized Inverse of Matrices and its Applications,* John Wiley, New York.

[8] V. Valerio (1976) – *Sulle inverse generalizzate e sulla risoluzione di particolari sistemi di equazioni lineari, con applicazione al calcolo delle strutture reticolari,* « Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. », *60,* 84–89.