PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série tome 45 (59), 1989, pp. 57-63

ON SOME GENERALIZED INVERSES OF MATRICES AND SOME LINEAR MATRIX EQUATIONS

Jovan D. Kečkić

Abstract. We point a generalized inverse A_G of a singular square complex matrix A, with the property that the general solution of the equation $A^n x = 0$ (and many other equations) can be expressed by means of A_G for all positive integers n. This inverse is a solution of the system (5), also satisfied by the strong spectral inverse of Greville. Applications to various matrix equations and to linear algebraic systems are given.

1. All matrices considered here will be square complex matrices. Although the theory of generalized inverses is well developed and pretty well known, in order to make this note self-contained, we begin by listing a few facts and definitions.

Suppose that A is a given matrix.

(i) Any matrix X satisfying the equality $AXA = A$ is called a (1)-inverse of A and is denoted by $A^{(1)}$. There is an infinity of such inverses.

(ii) There is a unique matrix X satisfying the equalities

$$AXA = A, \quad XAX = X, \quad (AX)* = AX \quad (XA)* = XA, \tag{1}$$

where $M*$ is the conjugate transpose of M. This unique solution of (1) is called the Moore-Penrose inverse of A and is denoted by A^+.

(iii) The smallest positive integer k such that rank A^{k+1} =rank A^k is called the index of A and is denoted by Ind$A$.

(iv) If $m(x) = x^s + c_{s-1}x^{s-1} + \cdots + c_k x^k$, with $c_k \neq 0$, is the minimal polynomial of A, then k =Ind$A$. The converse is also true.

(v) There is a unique matrix satisfying the equalities

$$AX = XA, \quad X^2 A = X, \quad A^{k+1}X = A^k \quad (k = \text{Ind}\, A). \tag{2}$$

This unique solution of (2) is called the Drazin inverse of A and is denoted by A^D.

AMS Subject Classification (1980): Primary 15 A 09, 15 A 24

The proofs of these and many other properties of various generalized inverses can be found e. g. in monographs [**1, 2, 3**]. We might add that statements (i) and (ii) remain valid even if A is a rectangular matrix.

2. "The basic application of inverses is for solving linear equations", says Ben-Israel [**4**], and indeed the first application given by Penrose [**5**] is contained in the following theorem.

THEOREM P. *A necessary and sufficient condition for the consistency of the equation* $AXB = C$ *is that* $AA^+CB^+B = C$, *in which case the general solution is* $X = A + CB^+ + T - A^+ATBB^+$, *where* T *is arbitrary.*

Consider now the equation

$$A^m X B^n = C \qquad (m, n \text{ positive integers}). \tag{3}$$

Of course, Theorem P can be applied to this equation, and its solution can be expressed in terms of $(A^m)^+$ and $(B^n)^+$, but generally speaking not in terms of A^+ and B^+, since the equalities $(A^+)^n = (A^n)^+, (B^+)^m = (B^m)^+$ need not be true, nor is there a generally valid formula connecting $(A^n)^+$ and A^+ for $n = 1, 2, \dots$.

On the other hand, Penrose also noted that Theorem P remains valid if A^+, B^+ are replaced by any (1)-inverses $A^{(1)}, B^{(1)}$ of A and B. Therefore, if there existed (1)-inverses $A^{(1)}, B^{(1)}$ of A and B such that

$$(A^{(1)})^n = (A^n)^{(1)},\ (B^{(1)})^n = (B^n)^{(1)}\ (n = 1, 2, \dots)$$

then it would be possible to express the general solution of (3) in terms of those inverses only. In fact, such inverses do exist and they can be successfully applied not only to (3), but also to other matrix equations.

We note that the equation (3) was solved by Cline [**6**], but only in m is a multiple of $p \geq$Ind A, and if n is a multiple of $q \geq$Ind B. In order to do that he introduced the so-called left and right power inverses for a matrix M^K where $k \geq$Ind M. Clin's inverses are interesting and usuful, as shown e. g. in [**7**], but their application to the equation (3) is not. Namely, if $m \geq$Ind A and $n \geq$Ind B, then $(A^D)^m$ and $(A^D)^n$ are (1)-inverses of A and B, and hence the equation (3) can be solved in terms of the Drazin inverses of A and B.

3. Let Ind $A = k$. Instead of looking for matrices X such that

$$A^n X^n A^n = A^n \qquad (n = 1, 2, \dots), \tag{4}$$

we shall look for matrices X with the properties

$$A^k X = X A^k, \quad A^n X^n A^n = A^n \quad (n = 1, 2, \dots, k-1). \tag{5}$$

Of course, (4) is an easy consequence of (5); besides, inverses with the additional property that they commute with A^p for $p \geq k$ are particularly usuful in equation solving, as will be seen later.

Any matrix satisfying (5) will be denoted by A_G. The existence of such matrices follows from the existence of the strong spectral inverse, introduced by Greville [**8**], which will be mentioned later. Nevertheless, we indicate a method for obtaining some (but not all) generalized inverses A_G.

Let $J = P^{-1}AP$, where P is nonsingular, be the Jordan canonical form of A. Then

$$J = J_1 \oplus J_2 \oplus \cdots \oplus J_s \oplus R, \tag{6}$$

where the blocks $J_1, \ldots, J_s$ are nilpotent (the only possible nonzero entries being l's immediately above the diagonal) and R is nonsingular. In other words, each of the nilpotent blocks is either a zero matrix or has the form

$$B = \begin{Vmatrix} O_{r-1,1} & I_{r-1} \\ O_{1,1} & O_{1,r-1} \end{Vmatrix} \quad (1 \le r \le k),$$

where $O_{p,q}$ is the zero matrix with p rows and q columns, and I_p is the unit matrix of order p. Any (1)-inverse od B has the form

$$B^{(1)} = \begin{Vmatrix} C & D \\ I_{r-1} & E \end{Vmatrix},$$

where C, D, E are arbitrary matrices of correct size, and it is easily verified that $B^kB^{(1)} = B^{(1)}B^k, B^n(B^{(1)})^nB^n = B^n$ for $n = 1, 2, \ldots, k-1$. Hence, if $A = PJP^{-1}$, and if J is given by (6), any matrix

$$A_G = P(J_1^{(1)} \oplus J_2^{(1)} \oplus \cdots \oplus J_s^{(1)} \oplus R^{-1})P^{-1} \tag{7}$$

satisfies the conditions (5).

We list a few properties of A_G which will be used later:

(i) $A^nA_G = A_GA^n$ for all $n \ge$Ind A;

(ii) $A^nA_G^nA^p - A^p$ if $p \ge n$;

(iii) $A^{n+p}A_G^n = A^p$ for all $n = 1, 2, \ldots$ and all $p \ge$Ind A.

The first two are direct consequences of (5), and we only prove the third.

Let m be the smallest positive integer such that $mp \ge n$. Since $p \ge$Ind A, we have

$$\begin{aligned} A^{n+p}A_G^n &= A^nA_G^nA^p = A^nA_G^pA^pA_G^{n-p} = A^nA_G^{n-p} = \\ &= A^{n-p}A_G^{n-p}A^p = \cdots = A^{n-(m-1)p}A_G^{n-(m-1)p}A^p = A^p, \end{aligned}$$

by (ii), as $p \ge n - (m-1)p$.

4. Using the inverses introduced in the previous section we can solve the equation (3) by a direct application of Theorem P. Indeed, (3) is consistent if and only if

$$A^mA_G^mCB_G^nB^n = C, \tag{8}$$

in which case its general solution is $X = A_G^mCB_G^n + T - A_G^mA^mTB^nB_G^n, T$ arbitrary.

A special case of (3) is interesting.

THEOREM 1. *If m, n, p are nonnegative integers, the equation*

$$A^m X A^n = A^p \tag{9}$$

is consistent if and only if $p \geq \max(m,n)$ or $p \geq$Ind A.

Proof. If $p \geq \max(m,n)$, the consistency criterion is fulfilled, since $A^m A_G^m A^p A_G^n A^n = A^p A_G^n A^n = A^p$. Let $\max(m,n) > p \geq$Ind A. Then

$$A^m A_G^m A^p A_G^n A^n = A^{m+p} A_G^m A_G^p A^n = A^p A_G^n A^n = A_G^n A^{n+p} = A^p.$$

It remains to prove that

$$p < \max(m,n) \quad \text{and} \quad p < \operatorname{Ind} A \tag{10}$$

implies that (9) is inconsistent. Let $k =$Ind A. Then the minimal polynomial of A has the form

$$(t^s + c_{s-1}t^{s-1} + \cdots + c_0)t^k \quad \text{with} \quad c_0 \neq 0, \tag{11}$$

and hence

$$(A^s + c_{s-1}A^{s-1} + \cdots + c_0 I)A^k = 0 \ \ (c_0 \neq 0). \tag{12}$$

Suppose that (10) is true and let $\max(m,n) = m$. If $k \leq m$, the consistency criterion (8), with $C = A^p$, cannot be fulfilled, since (12) after postmultiplication by $A^{m-k} A_G^m A^p A_G^n A^n$ would imply that $(A^s + c_{s-1}A^{s-1} + \cdots + c_0 I)A^p = 0$, $c_0 \neq 0$; i. e. that (11) is not the minimal polynomial of A. On the other hand, if $k > m$, then $A^m A_G^m A^p A_G^n A^n = A^p$ woul imply $A^k A_G^m A^p A_G^n A^n = A^{k-m+p}$, and (12) after postmultiplication by $A_G^m A^p A_G^n A^n$ would become $(A^s + c_{s-1}A^{s-1} + \cdots + c_0 I)A^{k-m+p} = 0$. Since $k - m + p < k$, this would again mean that (11) is not the minimal polynomial of A.

If $\max(m,n) = n$, similar contadictions are obtained if (12) is premultiplied by $A^m A_G^m A^p A_G^n A^{n-k}$ (if $k \leq n$), or by $A^m A_G^m A^p A_G^n$ (if $k > n$).

The proof is complete.

The equation (9) can be generalized, since the following result is true.

THEOREM 2. *The equation $A^m X A^n = a_0 I + a_1 A + \cdots + a_p A^p$ is consistent if and only if for every $i = 1, 2, \ldots, p$ we have: $a_i \neq 0$ implies that $A^m X A^n = A^i$ is consistent.*

Remark. The above method can be extended to handle some (but not all) equations of the form $A^m X A^n = A^p B^q$. Without ging into details, we give an example: If Ind $A \leq 3$, the equation $A^5 X B^3 = A^3 B^4$ is consistent.

We shall now consider two systems of equations.

THEOREM 3. *Suppose that m and n are positive integers such that $n \geq m$. The system*

$$AXA^m = A^m, \qquad A^n X = XA^n \tag{13}$$

is consistent if and only if $n \geq$Ind A, *in which case its general solution is*

$$X = A_G + T - A_G^n A^n T + A_G^n A^n T A^m A_G^m - T A^n A_G^n + A_G A T A^n A_G^n - A_G A T A^m A_G^m, \tag{14}$$

where T *is arbitrary.*

Proof. If $n \geq$Ind A, the system (13) is consistent, since $X = A_G$ solves it. Conversely, suppose that $k =$Ind $A > n \geq m$ and that (13) is consistent. The minimal polynomial of A has the form (11) and hence the equality (12) is true. But from (13) follows $A^k X = A^{k-1}$ and postmultiplying (12) by X we conclude that (11) is not the minimal polynomial of A.

In order to obtain the general solution of (13) we proceed as follows. The general solution of the first equation of (13) is

$$X = A_G + U - A_G A U A^m A_G^m \qquad (U \text{ arbitrary}) \tag{15}$$

Substituting (15) into (13) we obtain

$$A^n U - A^n U A^m A_G^m = U A^n - A_G A U A^n, \tag{16}$$

which implies, after postmultiplication by $A^m A_G^m$, that

$$(I - A_G A) U A^n = 0 \tag{17}$$

The general solution of (17) is given by

$$U + V - (I - A_G A) V A^n A_G^n \quad (V \text{ arbitrary}) \tag{18}$$

We now substitute (18) into (16) and we get

$$A^n V (I - A^m A_G^m) = 0.$$

The general solution of the last equation is

$$V = T - A_G^n A^n T (I - A^m A_G^m) \quad (T \text{ arbitrary}). \tag{19}$$

From (15), (18), (19) we conclude that (13) implies (14). Conversely, it is easily verified that (14) is a solution of (13) and the proof is complete.

Remark. For $m = n = 1$ we obtan the well-known result: There exists a commuting (1)-inverse of A if and only if Ind $A \leq 1$. For $m = 1$ we obtain the system considered and solved in [**9**] and [**10**].

The following result can also be proved by similar procedure.

THEOREM 4. *Suppose that* m, n *are positive integers and that* $n \geq m, n \geq$Ind A. *The general solution of the system*

$$A^m X A^m = A^M, \qquad A^n X = X A^n$$

is given by

$$X = A_G^m + T - A_G^n A^n T + A_G^n A^n T A^m A_G^m - T A^n A_G^n + A_G^m A^m T A^n A_G^n - A_G^m A^m T A^m A_G^m.$$

where T is arbitrary.

5. Regarding linear algebraic systems

$$A^n x = b \tag{20}$$

where n is a positive integer and x, b are column matrices, Cline [**6**] proved the following result.

THEOREM C *If $Ax = b$ is a consistent system of equations and if A and A^2 have the same rank, then the general solution of $A^n x = b, n = 1, 2, \cdots$, can be written as*

$$x = A_R^n b + (I - A_L A)y,$$

where y is arbitrary, and A_R, A_L are right and left power inverses of A.

Again we note that power inverses are not needed to solve the above system; it is easily solved by an application of the Drazin inverse, which is in the case Ind $A = 1$ (as implied by the conditions of Theorem C) called the group inverse and is denoted by $A^{\#}$. Indeed, if Ind $A = 1$, and if $Ax = b$ is consistent, the all the system $A^n x = b$ $(n = 1, 2, \dots)$ are consistent, and for any $n = 1, 2, \dots$ the general solution of (20) is

$$x = A^{\#} b + (I - A^{\#} A)y \quad (y \text{ arbitrary}).$$

It is not difficult to see that the group inverse $A^{\#}$, which is the unique solution of the equations

$$AXA = A, \quad AX = XA, \quad XAX = X, \tag{21}$$

can be replaced by any commuting (1)-inverse of A, i. e. by a matrix which satisfies only the first two equations of (21).

The inverses A_G enable us to solve (20) without the rather heavly restriction Ind $A = 1$.

THEOREM 5. *The system* (20) *is consistent if and only if $A^n A_G^n b = b$, and its general solution is then*

$$x = A_G^n b + (I - A_G^n A^n)y \quad (y \textit{ arbitraty}).$$

Let Ind $A = k$. Since $A^n A_G^n = A_G^n A^n = A^k A_G^k$ for $n \geq k$, we conclude that:

(i) If $A^k x = b$ is consistent, then all systems $A^n x = b \quad (n \geq k)$ are consistent;

(ii) If $n \geq k$, the general solution of the consistent system $A^n x = b$ is $x = A_G^n b + (I - A_G^k A^l)y \quad (y \text{ arbitrary})$;

(iii) If $k = 1$, Theorem 5 reduces to the result cited above, since A_G reduces to a commuting (1)-inverse of A.

6. As we mentioned earlier, Greville [**8**] introduced strong spectal inverses A^S of A and showed that they can be defined as solutions of the system

$$\begin{matrix} AXA = A, & A^kX = XA^k, & XA^n = A^{n-1}X^nA^n \\ XAX = X, & X^k = AX^k, & X^nA = X^nA^nX^{n-1} \end{matrix} \quad (k = \text{Ind}A;\ n = 2, 3, \dots, k) \qquad (22)$$

Since the equalities (5) are straight forward consequences of (22) we see that the equations considered in previous section could have been solved in terms of A^S. However, the inverses A_G suffice for that purpose, and it is easier to compute an A_G (a solution of (5)) then an A^S (a solution of (22)), just as it is easier to compute an $A^{(1)}$ than A^+. Notice that if $J = P^{-1}AP$ is the Jordan canonical form of A, then $A^S = PJ^+P^{-1}$. In fact, A^S is a special case of A_G, obtained when $J_k^{(1)}$ are replaced by J_k^+ in the formula (7).

7. The class of strong spectral inverses can be defined as follows. For a given matrix A, let $P(A)$ be the set of all nonsingular matrices P such that $J = P^{-1}AP = N \oplus R$ is the Jordan canaonical form of A, where N is nilpotent and R is nonsingular. Then $S(A) = \{P(N \oplus R^{-1})P^{-1} \mid P \in P(A)\}$ is the class of all strong spectral inverses of A.

Definition. If the set $P(A)$ contains a unitary matrix, we say that $A \in PP$, or that A is a PP matrix.

The class of PP matrices is a generalization of the class of EP matrices; see, for example, [**3**].

It is easily verified that if A is a PP matrix, then A^+ is a strong spectral inverse of A, and hence

$$(A^+)^n = (A^n)^+ \quad \text{for all } n = 1, 2, \dots \qquad (23)$$

It was known that (23) holds for EP matrices, and also that there are matrices, not EP, such that (23) id true. It would be interesting to find whether there is a matrix A, not PP, such that (23) remains true.

Finally, note that if A is a PP matrix, then the equations considered in Sections 4 and 5 can be solved in terms of A^+, i e. if A_G is replaced by A^+.

8. The author to express his appreciation to M. S. Stanković for many helpful discussions on generalized inverses and their applications.

REFERENCES

[1] C. R. Rao, S. K. Mitra, *Generalized Inverse of Matrices and its Applications*, New York-London-Sydney-Toronto, 1971.

[2] A. Ben-Israel, T. N. E. Greville, *Generalized Inverses: Theory and Applications*, New York-London-Sydney-Toronto, 1974.

[3] S. L. Campbell, C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, London-San Francisco-Melbourne, 1979.

[4] A. Ben-Israel, *Generalized inverses of matrices: a perspective of the work of Penrose*, Math. Proc. Cambridge Phil. Soc. **100** (1986), 407–425.

[5] R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Phil. Soc. **51** (1955), 406–413.

[6] R. E. Cline, *Inverses of rank invariant powers of a matrix*, SIAM J. Numer. Anal. **5** (1968), 182–197.

[7] T. N. E. Greville, *Some new generalized inverses with spectral properties*, p.p 26–46 in T. L. Boullion, P. L. Odell, *Proceedings of the Symposium on Theory and Applications of Generalized Inverses of Matrices*, Lubbock, Texas 1968.

[8] T. N. E. Greville, *Spectral generalized inverses of square matrices*, MRC Technical Summary Report 823, Mathematics Research Center, University of Wisconsin, Madison, Wisc. Oct. 1967.

[9] Л. Д. Добряков, *Коммутитующие обобщенные обратые матрицы*, Мат. Заметки **36** (1984), 17–23.

[10] J. D. Kečkić, *Commutative weak generalized inverses of a square matrix and some related matrix equations*, Pul. Inst. Math. (Beograd) **38(52)** (1985), 39–44.

Farmaceutski fakultet
11000 Beograd

(Received 08 10 1987)