A NEW CLASS OF g-INVERSE OF SQUARE MATRICES

By SUJIT KUMAR MITRA

Indian Statistical Institute

SUMMARY. Necessary and sufficient conditions are obtained for a matrix A to have a g-inverse with rows and columns belonging to specified linear manifolds. For a square matrix A, a g-inverse, with columns belonging to the linear manifold generated by the columns of A, is denoted by A_C^-. Such a g-inverse exists if and only if $R(A) = R(A^2)$. The following properties of A_C^- are established : (a) $A_C^- = A(A^2)^-$. (b) For any positive integer m, $(A_C^-)^m$ provides a reflexive g-inverse of A^m. (c) If x be an eigen vector corresponding to a nonnull eigen value λ of A, x is also an eigen vector of A_C^- corresponding to its eigen value $1/\lambda$. The converse of this result is also true. (d) A special choice of $(A^2)^- = (A^3)^-A$ leads to $A_C^- = A(A^3)^-A$ which is unique irrespective of the choice of $(A^3)^-$ and is in fact same as the Scroggs-Odell pseudoinverse (*J. SIAM* 1966) of A. When $R(A) = R(A^2)$ this indeed is a much simpler way of calculating the Scroggs-Odell pseudoinverse compared to the method indicated by its authors. (e) $A(A^3)^-A$ belongs to the subalgebra generated by A.

1. Notations

In this paper we consider matrices defined over the complex field. The following notations will be used throughout the paper. For a matrix A,

$R(A)$ represents the rank of A,

$\mathcal{M}(A)$ represents the linear space generated by the columns of A,

A^- represents a g-inverse as defined by Rao (1962, 1967), and

A_r^- is a reflexive g-inverse of A as defined by Rao (1967),

A_C^- is a symbol introduced in this paper to indicate a g-inverse of A whose columns are in $\mathcal{M}(A)$.

A_R^- indicates a g-inverse of A whose rows belong to the linear manifold generated by rows of A.

If B be a matrix,

$B \subset \mathcal{M}(A)$ indicates that columns of B belong to $\mathcal{M}(A)$. The symbol $\subset$ in other cases is used to denote set inclusion.

2. A class of g-inverse of square matrices

While looking for a g-inverse of a matrix A one sometimes enquires if A has a g-inverse G with columns belonging to a specified linear manifold $\mathcal{M}(P)$, and rows belonging to a manifold $\mathcal{M}(Q')$, that is a G of the form $G = PCQ$. The answer to this is contained in the following theorem.

Theorem 2.1 : *Given matrices P and Q, a necessary and sufficient condition for A to have a g-inverse of the form $G = PCQ$ is that*

$$R(QAP) = R(A). \quad \dots \quad (2.1)$$

Proof : Necessity follows from the definition of a g-inverse as given by Rao (1967) since, if $G = PCQ$ be a g-inverse of A, $A = AGA = AGAGA$. Hence $R(A) = R(AGAGA) = R(APCQAPCQA)$. This implies $R(QAP) \geqslant R(A)$. Trivially $R(QAP) \leqslant R(A)$. Hence $R(QAP) = R(A)$.

To prove sufficiency make repeated use of Corollary 1a.3 of Mitra (1968) to check that under condition (2.1) $P(QAP)^{-}Q$ is a g-inverse of A. In fact any g-inverse of the form $G = PCQ$ is always expressible as $P(QAP)^{-}Q$. For

$$APCQA = A \Longrightarrow QAPCQAP = QAP \Longrightarrow C = (QAP)^{-}.$$

Corollary 2.1 : *Given a matrix P, a necessary and sufficient condition for A to have a g-inverse of the form $G = PC$ is that*

$$R(AP) = R(A). \quad \dots \quad (2.2)$$

Proof : Take $Q = I$ in Theorem 2.1.

Of special interest is the case where A is square and P is A. By Corollary 2.1, A has a g-inverse belonging to $\mathscr{M}(A)$ if and only if $R(A^2) = R(A)$. For the convenience of future reference such a g-inverse whenever it exists will be indicated by $A_{\bar{C}}$.

3. Reflexivity and a stronger property

Obviously $R(A_{\bar{C}}) = R(A)$. Hence by Theorem 2a of Mitra (1968) $A_{\bar{C}}$ is a reflexive inverse of A, that is

$$A_{\bar{C}}AA_{\bar{C}} = A_{\bar{C}}, \qquad AA_{\bar{C}}A = A. \quad \dots \quad (3.1)$$

The following theorems present some useful properties of $A_{\bar{C}}$.

Theorem 3.1 : *For any positive integer m, $(A_{\bar{C}})^m$ is a reflexive inverse of A^m, that is, if $G = A_{\bar{C}}$,*

$$A^mG^mA^m = A^m \quad \dots \quad (3.2)$$

$$G^mA^mG^m = G^m \quad \dots \quad (3.3)$$

Proof : Since $G = AC$,

$$AG^2 = AGAC = AC = G. \quad \dots \quad (3.4)$$

(3.2) and (3.3) follow once the left hand side expressions are simplified by repeated use of (3.4), using (3.1) in the final step of simplification. A slightly more general version is given below.

Let $r_1, r_2, \ldots, r_n$ be a sequence of positive integers, $o = r_1+r_3+\ldots$, $e = r_2+r_4+\ldots$. The summation in o is taken over all r_i's in the sequence with an odd subscript i and similarly the summation in e is taken over all r_i in the sequence with an even subscript.

Theorem 3.2 (*A partial law of indices*) : *Consider the following statements. If n is odd and $o > e$*

$$A^{r_1}G^{r_2}A^{r_3}\ldots G^{r_{n-1}}A^{r_n} = A^{o-e} \quad \ldots \quad (3.5)$$

$$G^{r_1}A^{r_2}G^{r_3}\ldots A^{r_{n-1}}G^{r_n} = G^{o-e} \quad \ldots \quad (3.5')$$

If n is even and $o < e$

$$A^{r_1}G^{r_2}A^{r_3}\ldots A^{r_{n-1}}G^{r_n} = G^{e-o} \quad \ldots \quad (3.6)$$

$$G^{r_1}A^{r_2}G^{r_3}\ldots G^{r_{n-1}}A^{r_n} = A^{e-o}. \quad \ldots \quad (3.6')$$

(3.5), (3.5′), (3.6) and (3.6′) $\Longleftrightarrow G = A_{\bar{G}}$.

Proof : ($\Longleftarrow$ part). For $n = 2$, (3.6) follows from (3.4). To establish (3.6′) notice that $G \subset \mathscr{M}(A)$ and $R(G) = R(A)$ implies $A \subset \mathscr{M}(G)$. Hence from (3.1) we have as in (3.4).

$$GA^2 = A. \quad \ldots \quad (3.4')$$

(3.6′) for $n = 2$ follows from (3.4′).

Similarly repeated application of (3.4) and (3.4′) establishes (3.5) and (3.5′) for $n = 3$.

The general case both for even and odd n is proved by induction. The $\Longrightarrow$ part is trivial.

4. Eigen values and vectors

Theorem 4.1 : *If x be an eigen vector of A corresponding to a non-null eigen value λ, x is also an eigen vector of $A_{\bar{G}}$ corresponding to its eigen value $1/\lambda$. The converse of this result is also true.*

Proof : Use (3.4′) to note that $Ax = \lambda x \Longrightarrow G(\lambda^2 x) = \lambda x \Longrightarrow Gx = \lambda^{-1}x$. The converse similarly follows from (3.4)

g-inverses with this property are briefly discussed in Section 9 of Rao (1967) where a reference is made to the method of construction given by Scroggs and Odell (1966). Scroggs and Odell however virtually require the explicit reduction of A to its Jordan canonical form. Hence, their g-inverse is computationally more difficult compared to $A_{\bar{G}}$ which exists whenever $R(A^2) = R(A)$. In fact we can claim for $A_{\bar{G}}$

some more properties, in addition to what is conveyed by Theorem 4.1. Starting with an arbitrary vector x_1 let $x_1, x_2 = (A-\lambda I)x_1, \ldots, x_k = (A-\lambda I)^{k-1}x_1$ be a chain of generalised eigen vectors corresponding to the nonnull eigen value λ of A, where k is the least integer for which $(A-\lambda I)^k x_1 = 0$. Consider the subspace $\mathcal{M}(x_1, x_2, \ldots, x_k)$ spanned by $x_1, x_2, \ldots, x_k$.

$\mathcal{M}(x_1, x_2, \ldots, x_k)$ is obviously invariant under A, that is, $x \epsilon \mathcal{M}(x_1, x_2, \ldots, x_k) \Longrightarrow Ax \epsilon \mathcal{M}(x_1, x_2, \ldots, x_k)$. We prove

Theorem 4.2 : *$\mathcal{M}(x_1, x_2, \ldots, x_k)$ is invariant under $A_{\bar{C}}$.*

Proof : The proof consists in showing that $A_{\bar{C}} x_i \epsilon \mathcal{M}(x_1, x_2, \ldots, x_k)$ for each $i = 1, 2, \ldots, k$, for which one uses (3.4′) as in the proof of Theorem 4.1, noting that

$$Ax_i = x_{i+1} + \lambda x_i \text{ for } i = 1, 2, \ldots, (k-1) \text{ and } Ax_k = \lambda x_k. \qquad \text{Q.E.D.}$$

Consider now the Jordan canonical representation of matrix A

$$A = L \begin{bmatrix} C_1 & 0 & \ldots & 0 \\ 0 & C_2 & \ldots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \ldots & C_k \end{bmatrix} L^{-1}$$

where C_i is a lower Jordan matrix of order r_i corresponding to some eigen value λ_i of A. It is well known that C_i is nonsingular if $\lambda_i \neq 0$ and if $\lambda_i = 0$, $R(C_i^m) = \max(0, r_i - m)$ for every positive integer m. Hence if $R(A^2) = R(A)$ it is clear that each Jordan matrix corresponding to a zero eigen value of A is of rank 0, therefore of order 1, that is the Segre characteristic of A corresponding to a zero eigen value is (1, 1,...,1) implying thereby that the multiplicity of a zero root is equal to the nullity of A. Let us therefore write

$$A = L \begin{bmatrix} C_1 & 0 & \ldots & 0 & 0 \\ 0 & C_2 & \ldots & 0 & 0 \\ . & . & \ldots & . & . \\ 0 & 0 & \ldots & C_m & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix} L^{-1} \qquad \ldots \ (4.1)$$

where each C_i is now assumed to be nonsingular. Consider the corresponding partition of L as $(L_1 : L_2 : \cdots : L_m : L_{m+1})$ and of $(L^{-1})'$ as $(M_1 : M_2 : \ldots M_m : M_{m+1})$ and use (3.4′) to note that

$$AL_i = L_iC_i \Longrightarrow A_{\bar{C}} L_i C_i^2 = L_i C_i \Longrightarrow A_{\bar{C}} L_i = L_i C_i^{-1} \quad (i = 1, 2, \ldots, m).$$

Also $$M'_{m+1}AL = 0 \Longrightarrow M'_{m+1}A = 0 \Longrightarrow M'_{m+1}A_C^- = 0.$$

Hence

$$A_C^- = L \begin{bmatrix} C_1^{-1} & 0 & \dots & 0 & J_1 \\ 0 & C_2^{-1} & \dots & 0 & J_2 \\ . & . & \dots & . & . \\ 0 & 0 & \dots & C_m^{-1} & J_m \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} L^{-1} \qquad \dots \quad (4.2)$$

where J_j are certain matrices, possibly nonnull.

5. The Scroggs-Odell pseudoinverse

In this section we establish that the class of g-inverse introduced in Section 2 includes as a special case the Scroggs-Odell pseudoinverse. This indeed is suggested by (4.2) since A_C^- in this expression gives the Scroggs-Odell pseudoinverse if only the J_j matrices turn out to be null. We prove

Theorem 5.1 : *If $R(A^2) = R(A)$, $A(A^3)^-A$ is the Scroggs-Odell pseudoinverse of A.*

Proof : We need Lemma 5.1.

Lemma 5.1 : *$R(A) = R(A^2) \Longleftrightarrow R(A) = R(A^m)$ for every positive integer $m \geqslant 3$.*

Proof of Lemma 5.1 : $R(A) = R(A^2) \Longleftrightarrow \mathcal{M}(A) = \mathcal{M}(A^2)$. Let A^2x be any vector in $\mathcal{M}(A^2)$.

$Ax \in \mathcal{M}(A) \Longrightarrow Ax \in \mathcal{M}(A^2) \Longrightarrow Ax = A^2y \Longrightarrow A^2x = A^3y \Longrightarrow A^2x \in \mathcal{M}(A^3)$ $\Longrightarrow \mathcal{M}(A^2) \subset \mathcal{M}(A^3)$. Obviously $\mathcal{M}(A^3) \subset \mathcal{M}(A^2)$. Hence $\mathcal{M}(A^2) = \mathcal{M}(A^3)$. This argument carried step by step will show $\mathcal{M}(A^2) = \mathcal{M}(A^m) \Longrightarrow R(A^2) = R(A^m)$ for any positive integer $m \geqslant 3$ and hence establish Lemma 5.1 since the '$\Longleftarrow$' part of the lemma is trivially true.

Lemma 5.1 together with Theorem 2.1 shows that $A(A^3)^-A$ is indeed a g-inverse of A. Rest of the proof follows the same lines of argument as in the derivation of expression (4.2).

Observe that if $R(A) = R(A^2)$, $A_R^- = (A^2)^- A$ is a g-inverse of A which has properties parallel to that of $A_C^- = A(A^2)^-$. In particular, similar to (3.4) and (3.4') we have for A_R^- the following identities

$$A^2A_R^- = A, \quad (A_R^-)^2A = A. \qquad \dots \quad (5.1)$$

Since $A_{RC}^- = A(A^3)^-A$ is both A_R^- and A_C^- using (3.4) and (5.1) we have

$$A(A_{RC}^-)^2A = A_{RC}^-A = AA_{RC}^-. \qquad \dots \quad (5.2)$$

This shows A_{RC}^{-} commutes with A. We have thus arrived at a simpler proof of the following result due to Englefield (1966), the necessity part of which is trivially seen to be true.

Theorem 5.2 : *A necessary and sufficient condition for the existence of a commuting g-inverse of A is that*

$$R(A) = R(A^2).$$

It may be noted that A_{RC}^{-} is indeed the unique reflexive commuting inverse of A which Englefield denotes by A_R.

We shall now prove

Theorem 5.3 : *A_{RC}^{-} is a polynomial in A with scalar coefficients.*

Proof : Let the polynomial equation of minimum degree (p) satisfied by A be written in the form

$$A^r = a_{r+1}A^{r+1}+a_{r+2}A^{r+2}+\ldots++a_pA^p \qquad \ldots \quad (5.3)$$

where clearly $r \geqslant 1$. Multiplying both sides by $(A_{RC}^{-})^{r+1}$ we have

$$A_{RC}^{-} = a_{r+1}AA_{RC}^{-}+a_{r+2}A+\ldots+a_pA^{p-r-1}.$$

Hence

$$AA_{RC}^{-} = a_{r+1}A+a_{r+2}A^2+\ldots+a_pA^{p-r},$$

and

$$A_{RC}^{-} = (a_{r+2}+a_{r+1}^2)A+(a_{r+3}+a_{r+1}a_{r+2})A^2$$
$$+(a_p+a_{r+1}a_{p-1})A^{p-r-1}+a_pA^{p-r}. \qquad \ldots \quad (5.4)$$

This completes the proof of Theorem 5.3. Incidentally,

$$A = A(A_{RC}^{-})A$$

is a polynomial equation of degree $p-r+1$ satisfied by A, which contradicts our assumption regarding the degree of the minimum equation for A unless $r = 1$. Hence $r = 1$ and one may rewrite the minimum equation (5.3) in the form

$$A = A^2P(A) \qquad \ldots \quad (5.5)$$

where $P(A) = a_2I+a_3A+\ldots+a_pA^{p-2}$.

Theorem 2 of Pearl (1966) shows that the Moore-Penrose inverse A^+ when it commutes with A can be expressed as a polynomial in A with scalar coefficients. That Pearl's result follows as a corollary to Theorem 5.3 is clear once it is recognised that, under the assumed conditions, A^+ is identical with A^-_{RC}, the unique reflexive commuting g-inverse of A.

Let $S(A)$ denote the subalgebra generated by A. We shall prove

Theorem 5.4 : *A necessary and sufficient condition for $S(A)$ to contain a g-inverse of A is that*

$$R(A) = R(A^2).$$

Proof : Since each member of $S(A)$ commutes with A the necessity part is seen to follow from Theorem 5.2. Theorem 5.3 shows the condition is sufficient.

The following result is easy to establish.

Theorem 5.5 : *If $R(A) = R(A^2)$, each g-inverse in $S(A)$ can be expressed as*

$$A^-_{RC}+c[I-AP(A)]$$

where c is a scalar and $P(A)$ is as defined in (5.5).

6. Some remarks on conditions (3.2) and (3.3)

It is easily seen that $G = (A^2)^-A = A^-_R$ satisfies Theorem 3.1. It also satisfies Theorem 4.1 with A' and G' replacing A and G of the theorem. It seems therefore interesting to speculate if (3.2) and (3.3) holding for all positive integers implies that G is either A^-_R or A^-_C. The following theorem provides only a partial answer.

Theorem 6.1 : *If* (3.2) *holds good for all positive integers m, then the reciprocal of every nonnull eigen value of A is an eigen value of G.*

Proof : Let $G^k+a_1G^{k-1}+\ldots+a_kI$ be the minimum polynomial for G.

Then

$$A^k(G^k+a_1G^{k-1}+\ldots+a_kI)A^k = 0.$$

Hence using (3.2) we have

$$A^k+a_1A^{k+1}+\ldots+a_kA^{2k} = 0. \qquad \ldots \quad (6.1)$$

Let x be an eigen vector of A corresponding to its nonnull eigen value λ, then

$$\lambda^k(1+a_1\lambda+\ldots+a_k\lambda^k)x = 0 \qquad \ldots \quad (6.2)$$

which implies $1+a_1\lambda+\ldots+a_k\lambda^k = 0$ since $\lambda^k \neq 0$ and $x \neq 0$. Hence $\mu = 1/\lambda$ satisfies the equation

$$\mu^k+a_1\mu^{k-1}+\ldots+a_k = 0. \qquad \ldots \quad (6.3)$$

This suggests $1/\lambda$ is an eigen value of G. If (3.2) and (3.3) both hold good for all positive integers m, every nonnull eigen value of A is the reciprocal of an eigen value of G and vice versa.

REFERENCES

ENGLEFIELD, M. J. (1966) : The commuting inverses of a square matrix. *Proc. Camb. Phil. Soc.*, 62, 667-671.

MITRA, S. K. (1968) : On a generalised inverse of a matrix and applications. *Sankhya*, Series A, 30, 107-114

PEARL, M. H. (1966) : On generalised inverse of matrices. *Proc. Camb. Phil. Soc.*, 62, 673-677.

RAO, C. R. (1962) : A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *J. Roy. Stat. Soc.*, B, 24, 152-158.

——— (1967) : Calculus of generalized inverse of matrices, Part I : General theory. *Sankhyā*, Series A, 29, 317-342.

SCROGGS, J. E. and ODELL, P. L. (1966) : An alternative definition to the pseudoinverse of a matrix. *J. SIAM*, 14, 796-810.

Paper received : April, 1968.