ON A GENERALISED INVERSE OF A MATRIX AND APPLICATIONS

By SUJIT KUMAR MITRA

Indian Statistical Institute

SUMMARY. As a sequel to Rao (1967) this paper develops further the concept of generalised inverse (g-inverse) of a matrix. The class of g-inverses with the reflexive property is characterised. General solutions (X) are obtained for the matrix equations $XBX = X$ and $BXBXB = BXB$.

1. Definition of generalised inverse

We assume the reader is familiar with the notations in Rao (1965, 1967) which we adopt except for minor differences. For a matrix B, $\mathcal{M}(B)$ is used to denote the linear manifold generated by the columns of B. Orthogonal complement of $\mathcal{M}(B)$ is denoted by $\mathcal{O}(B)$, a null matrix by 0 and a null vector by $\mathbf{o}$.

Theorem 1a : *Let $B(m\times n)$ and $G(n\times m)$ be matrices such that*

$$H = GB \text{ is idempotent} \quad \ldots \quad (1.1)$$

$$\text{Rank } H = \text{Rank } B. \quad \ldots \quad (1.2)$$

Then (1.1) *and* (1.2) $\Longrightarrow BGB = B$, *i.e., G is a g-inverse of B in the sense of Rao* (1967), *denoted by B^-.*

Proof: It is obvious that $\mathcal{M}(H') \subseteq \mathcal{M}(B')$. (1.2) will therefore imply $\mathcal{M}(H') = \mathcal{M}(B')$. Hence $\mathcal{O}(H') = \mathcal{O}(B')$ and $H(I\text{-}H) = 0$ implies

$$B(I-H) = B(I-GB) = 0. \qquad \text{Q.E.D.}$$

Noting that the converse ('implied by' part) of Theorem 1a is contained in Theorem 2a of Rao (1967) one could put forward (1.1) and (1.2) for an alternative definition of a generalised inverse essentially equivalent to his definitions 1 and 2.

Corollary 1a.1 : *BG idempotent and Rank BG = Rank B imply and are implied by $G = B^-$.*

Corollary 1a.2 : *If GB is idempotent and Rank GB = Rank G or Rank B, then BG is idempotent. In the former case $G = B^-$ and in the latter case $B = G^-$.*

Corollary 1a.3 : *If $B = CD$ and Rank B = Rank D then $D^- = B^-C$ is a g-inverse of D and $D^-C^- = (CD)^-$ for any g-inverse C^- of C.*

A similar result holds when Rank B = Rank C.

Theorem 1b : *A necessary and sufficient condition for $G(n\times m)$ to be a g-inverse of $B(m\times n)$ is that*

$$\text{Rank } [I-GB] = n-\text{Rank } B \quad \ldots \quad (1.3)$$

Proof: Necessity is a consequence of the idempotency of matrix GB, implied by $G = B^-$. To prove sufficiency we need Lemma 1a.

Lemma 1a : *For a square matrix H of order n if*

$$Rank\ [I-H] = n - Rank\ H, \quad \dots \quad (1.4)$$

then H is idempotent.

Proof of Lemma : If $x \in \mathcal{C}(I-H')$, then

$$(I-H)x = o.$$

Hence $x = Hx$, which shows $\mathcal{M}(H) \supset \mathcal{C}(I-H')$. (1.4) will therefore imply $\mathcal{M}(H) = \mathcal{C}(I-H')$. Consequently

$$(I-H)H = 0.$$

This completes the proof of Lemma 1a.

Since $[I-GB]+GB = I$, clearly Rank $GB \geqslant n-$Rank $[I-GB]$. But Rank $GB \leqslant$ Rank B. Hence from (1.3), we have

$$\text{Rank}\ [I-GB] = n - \text{Rank}\ GB.$$

Theorem 1b follows from Lemma 1a and Theorem 1a. Note that (1.3) in Theorem 1b could be replaced by Rank $[I-BG] = m-$Rank B.

2. Reflexive type g-inverses

A reflexive type g-inverse of a matrix B (to be denoted by B_r^-) has been defined in Rao (1967) as a g-inverse G with the added property

$$B = G^-$$

We shall now prove the following theorem.

Theorem 2a : *A necessary and sufficient condition for a g-inverse G of B to be reflexive is that*

$$Rank\ G = Rank\ B. \quad \dots \quad (2.1)$$

Proof : Necessity is obvious. Sufficiency follows from Corollaries 1a.1 and 1a.2.

Frame (1964) uses the term semi-inverse to denote a g-inverse satisfying the rank condition (2.1). A semi-inverse is thus equivalent to a reflexive type g-inverse. Theorem 2a by itself is not useful in obtaining a reflexive type g-inverse since it does not say how to obtain a g-inverse with rank equal to the rank of the original matrix. For this we need the following characterisation of the class of reflexive type g-inverses.

Theorem 2b : *A necessary and sufficient condition for G to be a reflexive type g-inverse of B is that*

$$G = B^-BB^- \quad \dots \quad (2.2)$$

for some g-inverse B^- of B.

Proof : Necessity follows from the reflexive property

$$G = GBG.$$

To prove sufficiency note that $GB = B^-B$ is idempotent and of rank equal to rank of B. Hence by Theorem 1a, G is a g-inverse of B and Rank $G \geqslant$ Rank B. Trivially Rank $G \leqslant$ Rank B. Hence Rank $G =$ Rank B. Theorem 2b now follows from Theorem 2a.

Corollary 2b.1 : *Equation* (2.2) *could be replaced by*

$$G = B^-_{(1)} \, BB^-_{(2)} \qquad \dots \quad (2.3)$$

where $B^-_{(1)}$ *and* $B^-_{(2)}$ *are (possibly different) g-inverses of* B.

Corollary 2b.2 : B^- *in* (2.2) *replaced by* B^-_m (B^-_l) *gives a necessary and sufficient condition for* G *to be* $B^-_{mr}(B^-_{lr})$.

The following characterisation of a reflexive g-inverse is due to Frame (1964). The proof given below is however different and more direct.

Theorem 2c : *A necessary and sufficient condition for* G *to be a reflexive type g-inverse of* B *is that*

$$G = Q(PBQ)^+P \qquad \dots \quad (2.4)$$

where P *and* Q *are non-singular matrices and the superfix indicates a Moore-Penrose inverse.*

Proof : Theorem 2c can be deduced from the following results.

Lemma 2a : *If* H *be an idempotent matrix there exists a positive definite (p.d.) matrix* L *such that*

$$H^* = LHL^{-1}. \qquad \dots \quad (2.5)$$

Proof of Lemma 2a : If H is idempotent of order n and rank r, $I-H$ is also idempotent of same order and rank $n-r$. Consider the rank factorisations of H and $I-H$ as

$$H = MN, \quad I-H = M_1N_1.$$

Write $T_1 = (M : M_1)$, $T_2 = \begin{pmatrix} N \\ \cdots \\ N_1 \end{pmatrix}$, note that both T_1 and T_2 are square matrices of order n and

$$T_1 \, T_2 = I.$$

Hence
$$T_1 = T_2^{-1} = L^{-1}T_2^*, \text{ where } L = T_2^*T_2.$$

Therefore,
$$M = L^{-1}N^*.$$

Check
$$N^*M^* = LMNL^{-1}.$$

Lemma 2b : *If* $G(n \times m)$ *be a g-inverse of* $B(m \times n)$ *there exist p.d. matrices* U *and* V *of order* n *and* m *respectively such that*

$$(GB)^* = UGBU^{-1} \text{ and } (BG)^* = V^{-1}BGV. \qquad \dots \quad (2.6)$$

Lemma 2c : *For given p.d. matrices U and V the following sets of conditions on a matrix G are equivalent*

(i) $GBU^{-1}B^* = U^{-1}B^*$, $\quad G^* = V^{-1}BD$ *for some* D

(ii) $B^*V^{-1}BG = B^*V^{-1}$, $\quad G^* = DBU^{-1}$ *for some* D

(iii) $GBU^{-1}B^* = U^{-1}B$, $\quad BGVG^* = VG^*$

(iv) $B^*V^{-1}BG = B^*V^{-1}$, $\quad G^*UGB = G^*U$

(v) $BGB = B$, $GBG = G$, $(GB)^* = UGBU^{-1}$, $(BG)^* = V^{-1}BGV$.

Any one of (i)—(v) determines a unique g-inverse of B. Condition (v) has been proposed by Chipman (1964). Lemma 2c is similar to Theorem 5a and Lemma 5b of Rao (1967) and can be proved on similar lines.

Lemma 2d : *G as defined in Theorem 2c satisfies the conditions of Lemma 2c with $V = (P^*P)^{-1}$ and $U = (QQ^*)^{-1}$.*

3. The equation $XBX = X$

It will be seen that for a matrix $B(m \times n)$ the most general solution $X(n \times m)$ of equation

$$XBX = X \qquad \ldots \quad (3.1)$$

will also determine the entire class of matrices which have B for a generalised inverse.

Theorem 3a : *A general solution of equation* (3.1) *is*

$$X = C(DBC)_r^- D \qquad \ldots \quad (3.2)$$

where $C(n \times p)$ and $D(q \times m)$ are arbitrary matrices, and

$$\text{Rank } X = \text{Rank } DBC. \qquad \ldots \quad (3.3)$$

Proof: It is easy to verify that $C(DBC)_r^- B$ satisfies (3.1). Also, since an idempotent matrix is a reflexive g-inverse of itself, any X, for which (3.1) is true, can always be expressed as $X(BX)_r^-$ (or $(XB)_r^- X$). This shows that (3.2) with arbitrary choice of G and D indeed exhausts all solutions of (3.1). Observe

$$\text{Rank } X \leqslant \text{Rank } (DBC)_r^- = \text{Rank } DBC = \text{Rank } DBXBC \leqslant \text{Rank } X.$$

Corollary 3a.1 : *If Rank DBC = Rank D (or Rank C), (3.2) reduces to*

$$X = C(DBC)^- D. \qquad \ldots \quad (3.4)$$

Proof: Write $C(DBC)_r^- D = C(DBC)^- DBC(DBC)^- D$ and note that if Rank DBC = Rank D, by Corollary 1a.3 $BC(DBC)^- = D^-$. If Rank DBC = Rank C, $(DBC)^- DB = C^-$.

If X satisfies (3.1), $B = X^- = (XBX)^-$. Hence $X = X(XBX)^- X$. Obviously Rank XBX = Rank X. This shows that (3.4), together with Rank DBC = Rank D (or Rank C), C and D chosen arbitrarily otherwise, also gives a general solution of (3.1).

Corollary 3a.2 : *If Rank* $DBC = $ *Rank* $D =$ *Rank* C, $C(DBC)^- D$ *is unique with respect to the choice of the g-inverse of* DBC.

Other conditions that X may be required to satisfy in addition to (3.1) could be met by suitable choice of C, D and the g-inverse of DBC.

Examples : (i) $XBX = X, (BX)^* = BX$ or $B = X_m^-$.

In (3.2), choose D such that $D^*D = I(m\times m)$ and $DBC)_r^- = (DBC)_{lr}^-$. Note that since $F = DBC(DBC)_{lr}^-$ is hermitian, D^*FD is also hermitian.

(ii) $XBX = X$, $(XB)^* = XB$ or $B = X_l^-$.

In (3.2), choose C such that $CC^* = I(n\times n)$ and $(DBC)_r^- = (DBC)_{mr}^-$. Since $H = (DBC)_{mr}^- DBC$ is hermitian, CHC^* is also hermitian.

(iii) $XBX = X$, $BXB = B$ or $B = X_r^-$.

In (3.2), choose C and D such that Rank $DBC =$ Rank B. Use (3.3) and Theorem 2a.

(iv) $XBX = X$, $BXB = B$, $(BX)^* = BX$, $(XB)^* = XB$ or $B = X^+$ (the Moore-Penrose inverse).

Choose $C = D = B^*$. By Corollary 3a.1 we have

$$X = B^*(B^*BB^*)^- B^*.$$

From Theorem 5a(i) of Rao (1967) it follows that $X = B^+$.

(v) $XBX = X$, $BXB = B$, $(BX)^* = U(BX)U^{-1}$, $(XB)^* = V^{-1}(XB)V$ U and V are p.d. matrices (Chipman, 1964).

Factorise U and V as $U = P^*P$ and $V = QQ^*$. Choose $C = Q$, $D = P$ and $(DBC)_r^- = (DBC)^+$. We have thus

$$X = Q(PBQ)^+P.$$

Note that since $F = PBQ(PBQ)^+$ and $G = (PBQ)^+ PBQ$ are hermitian P^*FP $FP = U(BX)$ and $QGQ^* = (XB)V$ are both hermitian.

4. The equation $BXBXB = BXB$

The importance of this equation

$$BXBXB = BXB \qquad \dots (4.1)$$

is in its use by Khatri (1963) and Rao (1965), characterising quadratic forms in singular normal variables having a chisquare distribution. B in this context refers to the singular variance covariance matrix of the normal distribution and X the (symmetric) matrix of the quadratic form. It has been pointed out by Rao that in this context if B is further known to be nonsingular, for any solution X of (4.1), $B = X^-$. It seems now natural to enquire whether such a result would be true in general. We shall prove

Theorem 4a : *A necessary and sufficient condition for* (4.1) *to imply* $B = X^-$ *is that*

$$Rank\ BX = Rank\ XB = Rank\ X. \quad \ldots \quad (4.2)$$

Proof : Note that (4.2) implies

$$C(X'B') = C(X').$$

Hence
$$BX(I-BX)B = 0$$
implies
$$X(I-BX)B = 0.$$

Theorem 4a therefore follows from Corollary 1a.2. It is easy to construct examples to show that in (4.2), taken singly, Rank BX = Rank X or Rank XB = Rank X is not sufficient for Theorem 4a. Take for example

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and note that here CB is idempotent, though BC is not. These matrices satisfy $BCBCB = BCB$, Rank BC = Rank C and $CBCBC = CBC$, Rank BC = Rank B.

Corollary 4a.1 : *Condition* (4.2) *in Theorem* 4a *could be replaced by*

$$Rank\ BXB = Rank\ X. \quad \ldots \quad (4.3)$$

One may at this stage enquire if Rank BX = Rank XB = Rank B together with (4.1) would similarly imply $X = B^-$. The answer is in the negative as the following counter example will show. Take same B as before and

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Theorem 4b : *A necessary and sufficient condition for* (4.1) *to imply* $X = B^-$ *is that*

$$Rank\ BXB = Rank\ B \quad \ldots \quad (4.4)$$

The proof is similar to that of Theorem 1a.

Corollary 4b.1 : *If* B *factorises as* $B = B_2B_1$ *such that*

$$B_1XB_2 \text{ is idempotent}, \quad \ldots \quad (4.5)$$

then

$$tr\ BX = Rank\ B \text{ implies } (4.4). \quad \ldots \quad (4.6)$$

Proof : Note that

$$B_1XB_2B_1XB_2B_1XB_2 = B_1XB_2.$$

Hence Rank $BXB \geqslant$ Rank B_1XB_2 = Rank B as implied by (4.5) and (4.6). Trivially Rank $BXB \leqslant$ Rank B. Hence Rank BXB = Rank B.

Theorem 4c : *A general solution of equation* (4.1) *is*

$$X = C(DBC)_r^- D + E - B^-BEBB^- \quad \ldots \quad (4.7)$$

where $C(n \times p)$, $D(q \times m)$ *and* $E(n \times m)$ *are arbitrary matrices.*

Proof: We shall show that a general solution of (4.1) is

$$X = Y+W \quad \ldots \quad (4.8)$$

where Y and W are general solutions of $Y = YBY$ and $BWB = 0$ respectively.

It is easily verified that $Y+W$ satisfies (4.1). To show that $Y+W$ is in fact a general solution, it suffices to show that any solution of (4.1) can always be expressed in this form. For this, note that if X be a solution of (4.1), it also satisfies

$$B(X-XBXBX)B = 0, \quad \ldots \quad (4.9)$$

and

$$(XBXBX)B(XBXBX) = XBXBX. \quad \ldots \quad (4.10)$$

Theorem 4c follows from Theorem 3a and the general solution of equation $BWB = 0$, given by Rao (1967, Theorem 2d).

Corollary 4c.1: *If B is hermitian a general hermitian solution of equation* (4.1) *is*

$$X = CGCBCG^*C+E-HEH^* \quad \ldots \quad (4.11)$$

where C and E are arbitrary hermitian matrices,

$$G = (CBC)^- \text{ and } H = B^-B.$$

The type of factorisation visualised in the corollary is available in the situation considered by Rao (1965) and Khatri (1963). We have therefore the following result: 'A necessary and sufficient condition for a quadratic form (with coefficient matrix A) in singular normal variables (with variance covariance matrix Σ) to be distributed as a chisquare with d.f. equal to (a) Rank A or (b) Rank Σ is

$$\text{for (a)} : \Sigma = A^-$$

$$\text{for (b)} : A = \Sigma^-.$$

The same counter example as considered earlier in this section, with Σ substituted for B and A for X, shows that simple characterisation as in (a) or (b) above is not true always and a closer examination will reveal that the failure is entirely contributed by the indefiniteness of the matrix A of the quadratic form and the superfluity of its rank (in relation to the d.f. of chisquare), both permissible in this context, owing to the singularity of the variance covariance matrix. On the other hand, since the quadratic form has to be non-negative with probability one, there must exist a positive semidefinite quadratic form of rank equal to the d.f. of chisquare, such that the two forms differ only on a set of probability measure zero. The matrix B of the equivalent p.s.d. form is easily seen to be $B = A\Sigma A\Sigma A$. Observe that B satisfies

$$\Sigma B\Sigma B\Sigma = \Sigma B\Sigma \quad \ldots \quad (4.12)$$

$$\Sigma(A-B)\Sigma = 0 \quad \ldots \quad (4.13)$$

$$\Sigma = B^- \quad \ldots \quad (4.14)$$

Notice that $B = A\Sigma A$ satisfies (4.12) and (4.13) though not (4.14).

5. Concluding remarks

Alternative proofs of all theorems in this paper involving rank conditions could be obtained from the following theorem on rank of matrices. 'If P and Q be matrices of maximum possible rank, such that $PBC = 0$ and $ABQ = 0$, then Rank $ABC+$ Rank $B =$ Rank $AB+$Rank $BC+$ Rank PBQ (Khatri, 1961)'. We leave the details to the reader noting that the explicit use of linear manifolds is avoided in such a case.

The author is grateful to Dr. C. R. Rao and Dr. C. G. Khatri for valuable comments that have led to considerable improvement in the presentation of this paper.

References

Chipman, J. S. (1964): On least squares with insufficient observations. *J.A.S.A.*, 59, 1078–1111.

Frame, J. S. (1964): Matrix functions and applications—Part I. *IEEE Spectrum*, 1, No. 3, 208–220.

Khatri, C. G. (1961): A simplified approach to the derivation of the theorems on the rank of a matrix. *Journal of M.S.U. of Baroda*, 10, 1–5.

——— (1963): Further contributions to Wishartness and independence of second degree polynomials in normal vectors. *J.I.S.A.*, 1, 61–70.

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

——— (1965): *Linear Statistical Inference and its Applications*, John Wiley and Sons, New York and London.

——— (1967): Calculus of generalised inverse of matrices, Part I: General Theory. *Sankhyā*, Series A, 29, 317–342.

Paper received : August, 1967.