FURTHER CONTRIBUTIONS TO THE THEORY OF GENERALIZED INVERSE OF MATRICES AND ITS APPLICATIONS

By C. RADHAKRISHNA RAO

and

SUJIT KUMAR MITRA

Indian Statistical Institute

SUMMARY. This is a sequel to an earlier paper by the authors on the same subject presented at the Sixth Berkeley Symposium.

In the previous paper, the authors have discussed three basic types of g-inverses—the minimum norm g-inverse, the least squares g-inverse and the minimum norm least squares g-inverse. In this paper these concepts are extended to more general situations involving semi norms in place of norms used earlier.

It is shown that a matrix is uniquely determined by its class of g-inverses. Further the subclass of g-inverses with a specified rank is characterized. Partial isometries are discussed in a general set-up with reference to a pair of linear spaces furnished with arbitrary quadratic norms.

A unified theory of linear estimation is presented using the expression for a minimum semi norm inverse.

1. Introduction

This is a sequal to an earlier paper by the authors on the subject presented at the Sixth Berkeley Symposium in 1970. To make the discussion of the present paper self contained, the definitions, concepts and some basic theorems are restated.

In the previous paper, we have discussed the basic types of g-inverses, constrained inverses, and their applications in statistics and electrical network theory. In the present paper, we consider other types of g-inverses and partial isometry or semiunitary transformations.

Some of the notations used are explained in the text. The rest are standard notations (see e.g. Mitra and Rao, 1968).

Using the expression for a minimum semi norm inverse of a matrix, a general theory of least squares is developed where there is no need to examine for singularity or otherwise of the dispersion matrix of the observations. Expressions are given for variances and covariances of BLUE's and for test criteria. There is no need to examine the internal consistency of a linear hypothesis. It is automatically taken care of by a suitable test involving a g-inverse of a matrix.

2. Generalized inverse of a matrix

Definition 1 : Let A be an $m \times n$ matrix of arbitrary rank. A generalized inverse of A is an $n \times m$ matrix A^- such that A^-y is a solution of $Ax = y$ for any y which makes the equation consistent.

Definition 2 : A g-inverse of A of order $m\times n$ is a matrix A^- of order $n\times m$ such that

$$AA^-A = A. \qquad \ldots \quad (2.1)$$

Definition 3 : A g-inverse of A of order $m\times n$ is a matrix A^- of order $n\times m$ such that A^-A is idempotent and $R(A^-A) = R(A)$ or alternatively AA^- is idempotent and $R(AA^-) = R(A)$.

Theorem 2.1 :

(a) *A^- exists and $R(A^-) \geqslant R(A)$.*

(b) *If A^- is a specific g-inverse of A, the class of all g-inverses of A is given by*

(b1) $A^- + U - A^-AUAA^-$

where U is arbitrary, or equivalently,

(b2) $A^- + (I-A^-A)V + W(I-AA^-)$

where V and W are arbitrary matrices.

(c) *A matrix is uniquely determined by its class* (b1) *or* (b2) *of g-inverses.*

We prove (b2) by establishing its equivalence with (b1), for which note that (b1) is obtained from (b2) by choosing $W = U$ and $V = UAA^-$ and that (b2) is obtained from (b1) if we put $U = (I-A^-A)V + W(I-AA^-)$.

To prove (c) we show that if A and B are matrices such that $BA^-B = B$ for every g-inverse A^- of A, and $AB^-A = A$ for every g-inverse B^- of B then $A = B$.

Observe that $G = A^- + (I-A^-A)(I-A^-A)^*B^*$ is a g-inverse of A and

$$BA^-B = BGB = B \Longrightarrow B(I-A^-A) = 0. \qquad \ldots \quad (2.2)$$

Similarly it is seen that

$$(I-AA^-)B = 0 \qquad \ldots \quad (2.3)$$

thus implying $B = AWA$ for some matrix W, for example $W = A^-BA^-$. We note likewise that $A = BVB$ for some V. Then

$$AWA = B = BB^-B = AWAB^-AWA = AWAWA$$

$$\Longrightarrow A = BVB = AWAVB = AWAWAVB = AWA = B.$$

3. Solution of linear equations

Generalized inverses of matrices are useful in obtaining neat closed expressions for general solutions to consistent linear equations, as illustrated in the table below, where Z and z represent respectively an arbitrary matrix and an arbitrary vector of appropriate order, and G an arbitrary g-inverse of A.

equation	consistency condition	general solution
$Ax = y$	$AA^-y = y$	$x = A^-y+(I-A^-A)z$ or, if $y \neq 0$ $x = Gy$
$AXB = C$	$AA^-CB^-B = C$	$X = A^-CB^-+Z-A^-AZBB^-$
$AX = D,\ XB = E$	$AE = DB$	$X = A^-D+EB^--A^-AEB^- +(I-A^-A)Z(I-BB^-)$

The consistency conditions for all the three cases and the general solutions to equations $Ax = y$ and $AXB = C$ are obtained in Rao (1967).

If $y \neq 0$, choose V such that $Vy = z$ and write $G = A^-+(I-A^-A)V$. With such a choice of G the general solution $x = A^-y+(I-A^-A)z$ to equation $Ax = y$ is alternatively expressible as $x = Gy$. Check that G is also a g-inverse of A.

That $X = A^-D+EB^--A^-AEB^-+(I-A^-A)Z(I-BB^-)$ provides a general solution to equations $AX = D$, $XB = E$ is seen as follows. Clearly an X, so determined, satisfies $AX = D$, $XB = E$. Conversely, any X satisfying $AX = D$, $XB = E$ could be represented in this form, if Z is taken as $X-A^-D-EB^-+A^-AEBB^-$.

4. Projections

Theorem 4.1 :

(a) *A matrix P of order $m \times m$ represents a projection iff P is idempotent (i.e. $P^2 = P$) in which case it is a projection on $\mathcal{M}(P)$ along $\mathcal{M}(I-P)$.*

(b) *If for $x \epsilon E^m$, $y \epsilon E^m$ the inner product is defined to be $x^*\Lambda y$ where Λ is p.d., then*

(b1) *P represents an orthogonal projection iff*

$$P^2 = P \text{ and } \Lambda P = P^*\Lambda$$

(b2) *the orthogonal projection operator onto $\mathcal{M}(A)$ is given by*

$$P = A(A^*\Lambda A)^-A^*\Lambda.$$

Proof : (a) is proved in Rao (1967), (b1) and (b2) in Mitra and Rao (1968).

The following result is easy to prove.

Corollary : *Let $P^2 = P$. Then P represents an orthogonal projection if*

$$\Lambda = P^*P+(I-P)^*(I-P).$$

5. g-inverse of specified rank

Definition : G is said to be a reflexive g-inverse of A if

$$A = AGA \quad \text{and} \quad G = GAG$$

(i.e., $G = A^-$ and $A = G^-$ both hold true). A reflexive g-inverse of A is denoted by A_r^-. We have seen in Theorem 2.1(a) that $R(A^-) \geqslant R(A)$. The following theorem shows that the class of reflexive g-inverses of A is identical with the class of g-inverses with minimum permissible rank.

Theorem 5.1 : *A^- is a reflexive g-inverse of A iff any one of the following conditions hold*

(a) $R(A^-) = R(A)$

(b) $A^- = GAG$

for some g-inverse G of A.

For a proof of Theorem 5.1 see Mitra (1968a).

Theorem 5.2 : *Given g-inverses G_1 and G_2 of A, $A_r^- = G_1AG_2$ is the unique reflexive g-inverse of A such that $A_r^- A = G_1A$, $AA_r^- = AG_2$.*

Proof : $A_r^-A = G_1A,\ AA_r^- = AG_2 \Longrightarrow A_r^- = A_r^-AA_r^- = G_1AA_r^- = G_1AG_2$.

Theorem 5.3 : *Let A be a matrix of order $m\times n$ and rank r and s an integer, $r \leqslant s \leqslant \min(m, n)$. Then G is a g-inverse of A of rank s iff $G = (A+B)_r^-$ where B is such that*

$$R(A : B) = R(A^* : B^*) = R(A)+R(B) = s. \qquad \dots \quad (5.1)$$

Proof of 'if' part : Let (5.1) be true. Then

$$Aa = Bb \Longrightarrow Aa = Bb = 0. \qquad \dots \quad (5.2)$$

Further
$$R(A+B) = s \text{ and } \mathscr{M}(A) \subset \mathscr{M}(A+B) \qquad \dots \quad (5.3)$$

$$\begin{aligned}(5.3) &\Longrightarrow (A+B)\ (A+B)_r^-\ A = A \\ &\Longrightarrow A(A+B)_r^-\ A-A = -B(A+B)_r^-\ A \\ &\Longrightarrow A(A+B)_r^-\ A-A = 0\end{aligned}$$

in view of (5.2). Thus the 'if' part is established.

Proof of 'only if' part : Let G be A^- of rank of s. Consider $P = GA$ and Λ as in the corollary to Theorem 4.1. $P_G = G(G^*\Lambda G)^-G^*\Lambda$ is the orthogonal projector onto $\mathscr{M}(G)$ while P is the orthogonal projector onto $\mathscr{M}(GA) \subset \mathscr{M}(G)$. Thus P_G-P is the orthogonal projector onto the orthogonal complement of $\mathscr{M}(GA)$ in $\mathscr{M}(G)$ of dimension $s-r$. Let $Y = (G^*\Lambda G)^-G^*\Lambda-A$ and $B = YGY$.

Then $GY = P_G-P$. Hence $R(GY) = s-r$. Also $GB = GY \Longrightarrow R(B) = s-r$. Observe now that $G(A+B) = G(A+Y) = P_G \Longrightarrow R(A+B) = R(A)+R(B) = s$. Hence by definition 3 in Section 2, G is $(A+B)^-$, obviously $(A+B)_r^-$. This completes proof of 'only if' part.

6. Three basic types of inverses

In the previous paper we considered three basic types of g-inverses, depending on the nature of solution needed of a consistent or an inconsistent equation. We extend these types to more general situations involving semi norms in the place of norms used in the earlier work. We consider the semi norm $\|x\|_n = (x^*Nx)^{1/2}$ in E^n, where N is p.s.d., and the seminorm $\|y\|_m = (y^*My)^{1/2}$ in E^m where M is p.s.d. Let A be an $m \times n$ matrix.

Definition 1 : G is said to be semi norm g-inverse of A if for any y such that $Ax = y$ is consistent, $x = Gy$ is a solution with the least semi norm. In particular if the semi norm is defined by $\|x\|_n = (x^*Nx)^{1/2}$ the G inverse is represented by the symbol $A^-_{m(N)}$.

Theorem 6.1 : *If G is an N-semi norm g-inverse of A, then it is necessary and sufficient that*

$$AGA = A, \ (GA)^* N = NGA. \qquad \dots \quad (6.1)$$

Such a G exists and one choice of G is

$$(N+A^*A)^-A^*[A(N+A^*A)^-A^*]^- \qquad \dots \quad (6.2a)$$

and, in particular, when $\mathcal{M}(A^*) \subset \mathcal{M}(N)$, it has the form

$$N^-A^*(AN^- A^*)^-. \qquad \dots \quad (6.2b)$$

A general solution is

$$G+W(I-AG)+(I-GA)V \qquad \dots \quad (6.3)$$

where W is arbitrary and G is given by either (6.2a) *or* (6.2b) *as the case may be and V is an arbitrary solution of $N(I-GA)V = 0$.*

Proof : Conditions (6.1) are established in the same manner as in the proof of the corresponding results for $N = I$, given in Rao (1967).

To establish (6.2a) we observe that since N is n.n.d., $\mathcal{M}(A^*) \subset \mathcal{M}(N+A^*A)$ i.e. $A^* = (N+A^*A)B^*$ for some B. This implies

$$A(N+A^*A)^-A^* = B(N+A^*A)B^*$$

$$\Longrightarrow R[A(N+A^*A)^-A^*] = R[B(N+A^*A)B^*]$$

$$= R[B(N+A^*A)] = R(A).$$

Using definition 3 of Section 2 it is seen that G as given by (6.2a) is indeed a g-inverse of A. Further for this G

$$NGA = (N+A^*A)GA-A^*AGA$$

$$= A^*[A(N+A^*A)^-A^*]^-A-A^*A$$

$$\Longrightarrow NGA = (GA)^*N.$$

(6.2b) is obtained in a similar manner. (6.3) follows from the expression (6.2) for a general solution to a g-inverse of A since

$$N(I-A^-A)VA = A^*V^*(I-A^-A)^*N \Longrightarrow N(I-A^-A)VA = 0.$$

Corollary : *Some other choices of minimum N-seminorm g-inverse are as follows :*

(i) $(N+cA^*A)^-A^*[A(N+cA^*A)^-A^*]^-$

where c is any positive constant, and

(ii) $(N+W)^{-}A^{*}[A(N+W)^{-}A^{*}]^{-}$

where W is a n.n.d. matrix such that $\mathcal{M}(N)$ and $\mathcal{M}(W)$ are virtually disjoint and $\mathcal{M}(A^{*}) \subset \mathcal{M}(N : W)$. Note that (ii) *is a g-inverse of A but* (i) *is not necessarily so.*

Definition 2 : G is said to be a semi least squares g-inverse of A if the minimum of $\|y-Ax\|_m$ is attained at $x = Gy$ for any y. In partiular when $\|z\|_m = (z^{*}Mz)^{\frac{1}{2}}$, G will be referred to as M-semi least squares g-inverse of A and is denoted by $A^{-}_{l(M)}$.

Theorem 6.2 : *If G be a M-semi least squares g-inverse of A then it is necessary and sufficient that*

$$MAGA = MA, \ (AG)^{*}M = MAG. \qquad \dots \quad (6.4)$$

Such a G exists and a general solution to G is

$$(A^{*}MA)^{-}A^{*}M+[I-(A^{*}MA)^{-}A^{*}MA]U \qquad \dots \quad (6.5)$$

where U is arbitrary.

Proof : Conditions (6.4) are established in the same manner as in the proof of the corresponding results for $M = I$ given in Rao (1967). Proof of the rest of Theorem 6.2 is on the same lines as in Theorem 6.1 and is theorefore omitted.

Definition 3 : G is said to be semi norm semi least squares g-inverse of A if $x = Gy$ has minimum norm in the set of least squares solutions of $Ax = y$ for any y. In particular when the semi norms in E^{n} and E^{m} are defined by p.s.d. matrices N and M, then G is referred to as N-semi norm M-semi least squares g-inverse of A and denoted by A^{+}_{MN}.

Theorem 6.3 : *If G is N-semi norm M-semi least square inverse of A, then it is necessary and sufficient that*

$$MAGA = MA \qquad NGAG = NG$$

$$(AG)^{*}M = MAG, \ \ (GA)^{*}N = NGA. \qquad \dots \quad (6.6)$$

Proof : We omit the proof of Theorem 6.3 as it is similar to the proof of the corresponding proposition for $M = I$, $N = I$ given in Rao (1967).

Explicit expression for A^{+}_{MN} are found as follows.

The problem is to find $x = Gy$ such that $x^{*}Nx$ is a minimum subject to $A^{*}MAx = A^{*}My$. The equations are

$$Nx+A^{*}MA\lambda = 0 \qquad \dots \quad (6.7)$$

$$A^{*}MAx = A^{*}My \qquad \dots \quad (6.8)$$

which are obviously consistent. Let

$$\begin{pmatrix} N & A^{*}MA \\ A^{*}MA & 0 \end{pmatrix}^{-} = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}.$$

Then $x = C_2 A^* My$, in which case one choice of G (i.e. A^+_{MN}) is $G = C_2 A^* M$. Observe that such a matrix G is a g-inverse of A iff $R(A^*MA) = R(A)$. Explicit expressions for A^+_{MN} are given below :

Case 1. (N is p.d.) : Here it is easily seen that

$$A^+_{MN} = N^{-1}A^*MA(A^*MAN^{-1}A^*MA)^- A^*M.$$

This expression is unique and it does not depend on the choice of g-inverse involved.

Case 2. $[\mathcal{M}(A^*MA) \subset \mathcal{M}(N)]$: Here one choice of A^+_{MN} is given by

$$A^+_{MN} = N^- A^*MA(A^*MAN^- A^*MA)^- A^*M.$$

This expression is not unique as it may very well depend on the choice of N^-.

Case 3 (general case) : The general case can be brought under Case 2, if one considers instead of (6.7) and (6.8) the equivalent equations :

$$(N+A^*MA)x + A^*MA\lambda = A^*My \qquad \dots \quad (6.7')$$

$$A^*MAx = A^*My. \qquad \dots \quad (6.8')$$

Here one choice of A^+_{MN} is given by

$$G = (N+A^*MA)^- A^*MA[A^*MA(N+A^*MA)^- A^*MA]^- A^*M.$$

This expression is also not unique. In Case (2) the general solution to A^+_{MN} is

$$G + (I - GA)VA^*M$$

where G is a particular solution and V is an arbitiary solution of $N(I-GA)VA^*M = 0$.

7. Partial isometry (subunitary transformation)

Let us consider two finite dimensional vector spaces E^m and E^n furnished with inner products $(.,.)_m$, $(.,.)_n$ and associated norms $\|.\|_m$ and $\|.\|_n$. we define the adjoint $A^\#$ of an $m \times n$ matrix A, by the conditions

$$(Ax, y)_m = (x, A^\# y)_n \, \forall x \epsilon E^n, y \epsilon E^m. \qquad \dots \quad (7.1)$$

Let A be $m \times m$ (square matrix) such that

$$\|y_1 - y_2\|_m = \|Ay_1 - Ay_2\|_m \, \forall y_1, y_2 \epsilon E^m. \qquad \dots \quad (7.2)$$

Then it is known that $A^{-1} = A^\#$ and conversely. In such a case the square matrix A is said to be a unitary matrix (or unitary transformation). We shall extend this concept to linear transformations from E^n to E^m defined by an $m \times n$ matrix $A(y = Ax,\ x \epsilon E^n,\ y \epsilon E^m)$.

Definition[1] : An $m \times n$ matrix A is said to be a partial isometry (subunitary transformation) if

$$\|x_1 - x_2\|_n = \|Ax_1 - Ax_2\|_m \, \forall x_1, x_2 \epsilon \mathcal{M}(A^\#) \qquad \dots \quad (7.3)$$

which is equivalent to

$$(x_1, x_2)_n = (Ax_1, Ax_2)_m \, \forall x_1, x_2 \epsilon \mathcal{M}(A^\#) \qquad \dots \quad (7.4)$$

where $A^\#$ is the adjoint of A as defined in (7.1).

[1] A special case of this definition was considered by Erdelyi (1966), namely, when $(x_1, x_2)_n = x_2^* x_1$ and $(y_1, y_2)_m = y_2^* y_1$. Results proved here are parallel to those obtained by Erdelyi.

It may be noted that $\mathcal{M}(A^*)$ is same as the subspace in E^n which is orthogonal to null space of A (i.e. the set of all vectors x in E^n such that $Ax = 0$). It is clear that a relation such as (7.3) cannot hold for all $x_1, x_2 \epsilon E^n$ but only for a suitable subset. The concept of partial isometry is thus a natural generalization of a unitary transformation.

Theorem 7.1 : *A is a partial isometry iff*

$$AA^*AA^* = AA^* \quad \dots \quad (7.5)$$

or equivalently A^* is a Moore-Penrose inverse of A for appropriate norms in E^m and E^n.

Proof : Writing $x_1 = A^*y_1$ and $x_2 = A^*y_2$ and using (7.3) we have

$$(A^*y_1, A^*y_2) = (AA^*y_1, AA^*y_2) \forall y_1, y_2$$
$$\Longrightarrow (y_1, AA^*y_2) = (y_1, AA^*AA^*y_2) \forall y_1, y_2$$
$$\Longrightarrow AA^* = AA^*AA^*$$

which proves (7.5).

It easily follows from (7.5) that

$$AA^* = P_A \quad \text{and} \quad A^*A = P_{A^\#} \quad \dots \quad (7.6)$$

so that A^* the Moore-Penrose inverse.

In particular if the inner products in E^m and E^n are of the form $(x, y) = y^*x$, then a n.s. condition for A to be subunitary is

$$A^* = A^+. \quad \dots \quad (7.7)$$

Theorem 7.2 : *If A is a partial isometry $\|x\|_n = \|Ax\|_m$ if and only if $x \epsilon \mathcal{M}(A^*)$.*

Proof : The 'if part' follows from the definition of isometry. To prove the 'only if' part consider the orthogonal decomposition $x = x_1 + x_2$ where $x_1 \epsilon \mathcal{N}(A)$, the null space of A and $x_2 \epsilon \mathcal{M}(A^*)$. Then $Ax = Ax_2$ and $\|Ax\| = \|Ax_2\| = \|x_2\|$ since A is a partial isometry. But $\|Ax\| = \|x\| = (\|x_1\|^2 + \|x_2\|^2)^{1/2}$ then $\|x_1\| = 0$ or $x_1 = 0$.

Theorem 7.3 : *Let P and Q be orthogonal projectors of the same rank r onto subspaces of E^m and E^n respectively. Then there exists a partial isometry X of order $m \times n$ such that*

$$XX^* = P, \quad X^*X = Q, \quad PX = X, \; QX^* = X^*. \quad \dots \quad (7.8)$$

Proof : Note that P and Q are self adjoint idempotents in which case

$$P = LL^*, \quad Q = RR^* \quad \dots \quad (7.9)$$

where L is an $m \times r$ matrix and R is an $n \times r$ matrix, each of rank r such that $L^*L = R^*R = I_r$. Then it is easily seen that $X = LR^*$ satisfies (7.8) although the solution may not be unique and this is so whatever proper norm is taken in E^r for the definition of L^* and R^*. Indeed $X^* = X^+$ so that X is a partial isometry.

Remark : With L and R as defined in (7.9) the most general solution for the equations

$$XX^+ = P, \quad X^+X = Q, \quad PX = X, \; QX^* = X^*$$

is given by

$$X = LER^* \quad \dots \quad (7.10)$$

where E is an arbitrary nonsingular matrix.

Semiunitary matrix and supplements of a partial isometry. A partial isometry of full rank is called semiunitary. Notice that if the matrix A of order $m \times n$ is semiunitary, then either $AA^* = I_m$ or $A^*A = I_n$ according as min (m, n) is m or n.

Theorem 7.4 : *If A is a partial isometry (subunitary matrix) of order $m \times n$ and rank $r < \min(m, n)$, then there exists a partial isometry of the same order such that $A+B$ is semiunitary and $AB^* = 0$ or $B^*A = 0$ according as min (m, n) is m or n (such a matrix is called a supplement of the partial isometry A).*

Proof: Suppose $m = \min(m, n)$. Write $P = I_m - AA^*$ and let Q be an orthogonal projector of order $n \times n$ and rank $n-r$ such that $AQ = 0$. Determine B as in Theorem 7.3 to satisfy

$$BB^* = P, \qquad B^*B = Q, \qquad PB = B, \qquad QB^* = B^*$$

$$AQ = 0 \Longrightarrow AB^* = 0 \Longrightarrow (A+B)(A+B)^* = I.$$

Hence $A+B$ is semiunitary. The case where $n = \min(m, n)$ can be proved on similar lines.

The following theorem gives an interesting characterization of partial isometries in terms of semiunitary matrices

Theorem 7.5 : *Let A be a matrix of order $m \times n$ $(m < n)$. Then A is a partial isometry iff $A = UQ$ where U is semiunitary and Q is the orthogonal projector onto the range of A^*.*

Proof: The 'if' part is trivial. To prove the 'only if' part assume that A is a partial isometry, obtain B as in Theorem 7.4 and check that

$$A = (A+B)A^*A.$$

Theorem 7.6 : *The eigen values of a square partial isometry have absolute magnitudes on the closed interval* [0, 1].

Proof: Let A be a partial isometry and x an eigen vector of A corresponding to its eigen value λ, then

$$\lambda\bar{\lambda}(x, x) = (Ax, Ax) = (x, A^*Ax) = (x, Px)$$
$$= (x, P^*Px) = (Px, Px)$$

where $P = A^*A$ is the orthogonal projector onto $\mathcal{M}(A^*)$.

Hence $$|\lambda| \, \|x\| = \|Px\| \Longrightarrow |\lambda| = \frac{\|Px\|}{\|x\|} \leq 1.$$

Theorem 7.7 : *Let λ, x be the eigen value and the corresponding eigen vector of a partial isometry of order $n \times n$. Then*

(a) $\lambda = 0$ *iff* $x \in \mathcal{N}(A)$

(b) $|\lambda| = 1$ *iff* $x \in [\mathcal{N}(A)]^\perp$

(c) $0 < \lambda < 1$ *iff* $x \notin \mathcal{N}(A) \cup [\mathcal{N}(A)]^\perp$.

where $\mathcal{N}(B)$ denotes the null space of B.

Proof: (a) is easily proved. (b) $\Longleftrightarrow A^*Ax = x \Longleftrightarrow x \epsilon \mathcal{M}(A^*) = [\mathcal{N}(A)]^\perp$. (c) follows from (a), (b) and Theorem 7.6.

Corollary : *If the partial isometry A is* EPr *and λ is a nonnull eigen value of A, then $|\lambda| = 1$.*

Proof: Observe that the corresponding eigen vector $x \epsilon \mathcal{M}(A) = \mathcal{M}(A^*)^\perp$ since A is EPr. Apply Theorem 7.7.

Theorem 7.8 : *Let A be a partial isometry of order $m \times n$ and P, Q be unitary matrices of order $m \times m$ and $n \times n$ respectively. Then $B = PAQ$ is a partial isometry.*

Proof: Theorem 7.8 follows from the identities

$$PP^* = P^*P = I_m, \qquad QQ^* = Q^*Q = I_n.$$

Theorem 7.9 : *Let A, B be partial isometries of order $m \times n$ and $n \times p$ respectively, then $C = AB$ is a partial isometry iff*

$$A^*AB = B. \qquad \dots \quad (7.11)$$

Proof: $(7.11) \Longrightarrow (ABx,\ ABx) = (Bx,\ A^*ABx)$

$$= (Bx,\ Bx) = (x, x) \forall x \epsilon \mathcal{M}(C^*) \subset \mathcal{M}(B^*).$$

Conversely, if $Bx \epsilon \mathcal{M}(A^*)$

$$(ABx,\ ABx) < (Bx, Bx) = (x, Px) = (Px,\ Px)$$

$$\leqslant (x, x)$$

where P is the orthogonal projector B^*B.

8. Least squares theory with a possibly singular dispersion matrix

Let $Y = X\beta + \epsilon$ where $D(\epsilon) = \sigma^2 V$ be a Gauss-Markoff model with V possibly singular (p.s.d.). Let $p'\beta$ be an estimable parametric function, i.e., there exists a vector L such that $E(L'Y) \equiv p'\beta$ or $X'L = p$. In order to determine the BLUE of $p'\beta$ we need $L'Y$ such that $V(L'Y) = \sigma^2 L'VL$ is a minimum subject to $X'L = p$. The answer is indeed a minimum V-seminorm solution of $X'L = p$. Thus the optimum choice of L is

$$L = (X')^-_{m(V)} p = F'p \quad \text{say} \qquad \dots \quad (8.1)$$

and the BLUE is

$$p'FY. \qquad \dots \quad (8.2)$$

In Corollary 1 to Theorem 6.1 we have given two choices of minimum V-seminorm inverses. One choice is

$$(X')^-_{m(V)} = (V + cXX')^- X[X'(V + cXX')^- X]^- \qquad \dots \quad (8.3)$$

where c, in general, is any positive constant and may be zero when $\mathscr{M}(X) \subset \mathscr{M}(V)$. With such a choice as in (8.3), the BLUE is

$$p'[X'(V+cXX')^{-}X]^{-}X'(V+cXX')^{-}Y. \qquad \ldots \quad (8.4)$$

The form of expression (8.4) suggests that the BLUE of $p'\beta$ is $p'\hat{\beta}$ where $\hat{\beta}$ is the M-least squares solution of $Y = X\beta$, where

$$M = (V+cXX')^{-} \qquad \ldots \quad (8.5)$$

i.e., $\hat{\beta}$ is the value of β which minimises

$$(Y-X\beta)'(V+cXX')^{-}(Y-X\beta). \qquad \ldots \quad (8.6)$$

Thus we obtain a generalization of the least squares theory to any linear model *whether V is singular or not*. It may be noted that $(V+cXX')^{-}$ may not be a g-inverse of V, for which we could have also chosen M as

$$M = (V+W)^{-} \qquad \ldots \quad (8.7)$$

where W is a n.n.d. matrix such that $\mathscr{M}(V)$ and $\mathscr{M}(W)$ are virtually disjoint and $\mathscr{M}(X) \subset \mathscr{M}(V \vdots W) = \mathscr{M}(V+W)$. However the choice of M as in (8.5) is attractive as it involves directly the given matrices V and XX'. Because of the importance of the result (8.6) we state our unified theory of least squares as follows :

(i) Obtain $\hat{\beta}$ which minimises

$$(Y-X\beta)'(V+cXX')^{-}(Y-X\beta) \qquad \ldots \quad (8.8)$$

whether V is singular or nonsingular, where c can be any positive constant and zero when $\mathscr{M}(X) \subset \mathscr{M}(V)$. The normal equation leading to $\hat{\beta}$ is

$$X'(V+cXX')^{-}X\beta = X'(V+cXX')^{-}Y \qquad \ldots \quad (8.9)$$

giving a solution

$$\hat{\beta} = [X'(V+cXX')^{-}X]^{-}X'(V+cXX')^{-}Y. \qquad \ldots \quad (8.10)$$

(ii) The BLUE of $p'\beta$ is $p'\hat{\beta}$ and

$$V(p'\hat{\beta}) = \sigma^2 p'[(X'MX)^{-}-cI]p \qquad \ldots \quad (8.11)$$

$$\text{cov}(p'\hat{\beta}, q'\hat{\beta}) = \sigma^2 p'[(X'MX)^{-}-cI]q. \qquad \ldots \quad (8.12)$$

(iii) An unbiased estimator of σ^2 is

$$f^{-1}R_0^2 = f^{-1} \min_{\beta} (Y-X\beta)'(V+cXX')^{-}(Y-X\beta)$$

$$= f^{-1}[Y'(V+cXX')^{-}Y - Y'(V+cXX')^{-}X\hat{\beta}] \qquad \ldots \quad (8.13)$$

where

$$f = R(V \vdots X) - R(X). \qquad \ldots \quad (8.14)$$

It is interesting to see that the computed values of (8.11) and (8.12) are independent of c.

(iv) To test a set of k linear hypotheses, $p_i'\beta = d_i$, $i = 1, \ldots, k$ compute $u_i = p_i'\hat{\beta} - d_i$, the dispersion matrix $\sigma^2 D$ of $u' = (u_1, \ldots, u_k)$ by using the formulae (8.11, 8.12), and the statistics

$$T_1 = u'D^- u \quad \text{and} \quad T_2 = DD^- u - u \qquad \ldots \quad (8.15)$$

where D^- is any g-inverse of D. Let $h = R(D)$. The hypothesis is rejected if T_2 is nonnull or the statistic

$$F = \frac{u'D^- u}{h} \div \frac{R_0^2}{f} \qquad \ldots \quad (10.16)$$

as a variance ratio on h and f degrees of freedom exceeds some critical value.

Thus we have a simple and unified theory of least squares without having the need to examine whether V is singular or not. In the formulae (8.8) to (8.16) one can have any choice of the g-inverses involved.

REFERENCES

ERDELYI, I. (1966): On partial isometries in finite dimensional Euclidean spaces. *SIAM J. Appl. Math.* 14,, 453-467.

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

——— (1968b): A new class of g-inverse of square matrices. *Sankhyā*, Series A, 30, 323-330.

MITRA, S. K. and RAO, C. R. (1968): Some results in estimation and tests of linear hypotheses under the Gauss-Markoff model. *Sankhyā*, Series A, 30, 281-290.

RAO, C. R. (1967): Calculus of generalised inverse of matices, Part I: General theory. *Sankhyā*, Series A, 29, 317-341.

RAO, C. R. and MITRA, S. K. (1970): Generalized inverse of a matrix and its application, ISI-RTS Tech. Report No. 36/70 (To appear in the Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability).

Paper received : December, 1970.