In memory of my

brother and father

Ramasastry and Rama Rao

SOME CONTRIBUTIONS TO THE THEORY, APPLICATION AND COMPUTATION OF GENERALIZED INVERSES OF MATRICES

By

POCHIRAJU BHIMASANKARAM

A thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy

Calcutta

1971

ACKNOWLEDGEMENTS

I am very grateful to Professor C.R.Rao, F.R.S., Director of the Research and Training School, Indian Statistical Institute for having provided me the necessary facilities to carry out this work and for his many valuable suggestions and encouragement throughout this period. I express my deep sense of gratitude to Professor Sujit Kumar Mitra under whose supervision this work was carried out. I record my sincere thanks to him for his valuable guidance and encouragement which has been a rich source of inspiration for my research work. I also thank him for his kind permission to include our joint results in this thesis. I am highly grateful to Professor J. Roy for introducing me to the subject of generalized inverses of matrices and for his constant encouragement. How much I owe to these three Professors is known, perhaps, only to the four of us !

I am grateful to my collegues Mr. M.S.V.Rao and Mr. P.S.S.N.V.Prasada Rao for the many useful discussions I had with them on the contents of chapter 4.

Many faculty members of the Research and Training School have shown great interest in my work, to whom I am very thankful. I am thankful to the students of the Research and Training School, whose active and intelligent participation in my lectures on

Linear Algebra has considerably improved my appreciation of the subject.

I, finally, thank Mr. Arun Das for his neat and elegant typing of the thesis.

P.B.S.

CONTENTS

page

CHAPTER 2 (Contd.)

INTRODUCTION

The origin of the concept of a generalized inverse dates back to as early as 1920 when Moore defined the generalized inverse of a matrix which is equivalent to

Definition 1 (Moore) : Let A be a $m \times n$ matrix over the field of complex numbers. Then G is the generalized inverse of A if AG is the orthogonal projection operator projecting arbitrary vectors onto the column space of A and GA is the orthogonal projection operator projecting arbitrary vectors onto the column space of G.

Moore (1935) discussed this concept and its properties in some detail. Tseng (1949a, 1949b, 1956) discussed about generalized inverses of operators in more general spaces and Bjerhammer (1951) discussed the generalized inverse of a matrix in connection with an application to geodetic calculations.

Unaware of the work of Moore and others, Penrose (1955) defined a generalized inverse of a matrix as follows :

Definition 2 (Penrose) : Let A be a $m \times n$ matrix over the field of complex numbers. Then G is a generalized inverse of A if (i) $AGA = A$; (ii) $GAG = G$; (iii) $(AG)^* = AG$ and (iv) $(GA)^* = GA$.

Penrose (1955,1956) showed that for every matrix there exists a unique generalized inverse, discussed several of its important properties, gave applications to solution of matrix equations and suggested a practical method of computation of the generalised inverse.

As was pointed out by Rado (1956) Moore's definition of generalized inverse is equivalent to that of Penrose. Such generalized inverse is called the Moore-Penrose inverse and A^+ is used to denote the Moore-Penrose inverse of A.

Rao (1955), unaware of the earlier or contemporary work, constructed a pseudo-inverse of a matrix which he used in some least squares computations. In a paper in 1962, he defined a generalized inverse (g-inverse) as follows, proved some interesting properties and gave applications of g-inverses to Mathematical Statistics.

Definition 3 (Rao) : Let A be a $m \times n$ matrix. Then a $n \times m$ matrix G is a g-inverse of A if $x = Gy$ is a solution of the linear system $Ax = y$ whenever it is consistent.

Rao showed that Definition 3 is equivalent to

Definition 4 (Rao) : Let A be a $m \times n$ matrix. Then a $n \times m$ matrix G is a g-inverse of A if $AGA = A$.

A g-inverse of a matrix (in the sense of Rao) is in general not unique. As is easily observed (from definitions 2 and 4) the class of all g-inverses of a matrix A contains A^+. Rao (1965, 1967) developed a calculus of g-inverses, classified the g-inverses according to their use and according to the properties they possess similar to those of the inverse of a nonsingular matrix and suggested further applications to Mathematical Statistics. Mitra (1968a, 1968b) gave an equivalent definition of a g-inverse, developed further calculus of g-inverses, used g-inverses to solve some matrix equations of interest and explored the possibilities of some new classes of g-inverses with applications. In a series of papers, and a monograph Mitra and Rao (1968, 1970) pursued the research on generalized inverses of matrices and their applications to various scientific disciplines.

Some other principal contributors to the theory and application of generalized inverses are Ben Israel, Greville, Erdelyi, Odell, Bose and Khatri - to mention only a few. References to important contributions made by these people and others will be found in the monograph by Rao and Mitra (1970).

A g-inverse, in the sense of Rao, need only to satisfy a much weaker condition than the Moore-Penrose inverse and hence, in general, is much easier to compute. As has been shown by Mitra and Rao, for several applications Moore-Penrose inverse is not of paramount importance and any g-inverse satisfying weaker conditions would serve the purpose. In this thesis also, we point out a few such applications.

Each chapter of the thesis has a detailed introduction to it. Here we just mention briefly the problems considered.

We start with obtaining a characterisation of the Moore-Penrose inverse. In chapter 1, further, we discuss the inter-relationships of g-inverses of a matrix A and the powers of $A^{\#}A$, settle a conjecture of Mitra on g-inverses with specified manifolds, solve some matrix equations of interest suggesting an application of these to distribution of quadratic forms in normal variables and consider several other problems of interest in the calculus of g-inverses.

In chapter 2, g-inverses of $(A \vdots a)$ are obtained from those of A and vice versa (a is a column vector) and an application of these results to recalculation of least squares estimates for data or model changes is demonstrated. Also a result of Rohde (1965) on g-inverses of partitioned matrices is extended. It is shown that his result holds in more general set up than the one considered by him.

In chapter 3, the problem of simultaneous reduction of several hermitian forms is considered. We obtain solutions to this problem in several cases. A characterisation of semi-simple matrices with real eigen values is also given.

Chapter 4 is devoted to the computations of generalized inverses. Here two algorithms are suggested one of which is the usual Gaussian Elimination type algorithm useful in computing g-inverses of simple matrices and the other using Householder's transformations. The second, we believe, will be useful in computing g-inverses fairly efficiently. Two examples, worked out using the above algorithms, are also presented.

The following notations are used in the thesis. Matrices are denoted by capital letters A, B, Σ etc. and vectors by lower case letters x, y etc. In chapters 1,2 and 3 we consider matrices and vectors with elements defined over the field of complex numbers. In chapter 4 and section 2.6 we consider matrices over the field of real numbers. However, extensions of these results to the complex case are straightforward. Null matrix is denoted by 0. The symbols $\forall$, $\exists$ and $\ni$ denote 'for all', 'there exists' and 'such that' respectively. E^n denotes the n dimensional Unitary (or Euclidean) space.

Let $A = (a_{ij})$ be $m \times n$ matrix. Some functions of A and the symbols used are described in Table 1.

Table 1

Function	Symbol	Description
Transpose	A'	matrix with (i, j)th element $= a_{ji}$
Conjugate transpose	A^*	matrix with (i, j)th element $= \bar{a}_{ji}$
Rank	$R(A)$	the number of independent rows or columns of A.
Trace	tr A	$\Sigma\ a_{ii}$
Column space	$\mathcal{M}(A)$	vector space generated by the columns of A
Orthogonal complement	$A^{\perp}$	$\mathcal{O}(A) = \mathcal{M}(A^{\perp})$ is orthogonal complement of $\mathcal{M}(A)$
Adjoint	$A^{\#}$	matrix such that $(x, Ay) = (A^{\#} x, y)\ \forall\ x, y$ where $(.,.)$ is inner product.
Determinant	$\|A\|$	

Definitions of special matrices are given in Table 2.

Table 2

Type of matrix A	Definition
1. Symmetric	$A = A'$
2. Hermitian	$A = A^*$
3. Idempotent	$A^2 = A$
4. Positive definite (p.d.)	$x'Ax > 0$ for all x
5. Positive semi-definite(p.s.d.)	$x'Ax \geq 0$ for all x $x'Ax = 0$ for some x
6. Non-negative definite(n.n.d.)	$x'Ax \geq 0$ for all x
7. Semi-simple	$A = P \Lambda P^{-1}$ for some P and Λ diagonal
8. Orthogonal	$AA' = A'A = I$
9. Unitary	$AA^* = A^*A = I$

A classification of g-inverses is given in Tables 3 and 4. In the conditions G denotes a g-inverse of A of order $m \times n$ and rank r. P_X denotes the orthogonal projection operator onto $\mathcal{M}(X)$. For a matrix A, $\{A^-\}$, $\{A_r^-\}$, $\{A_\ell^-\}$ and $\{A_m^-\}$ denote respectively the class of all g-inverses, the class of reflexive g-inverses, the class of least squares g-inverses and the class of minimum norm g-inverses respectively.

Table 3 : Various types of g-inverses

Notation		equivalent conditions	purpose
A_L^{-1}		$GA = I$	solving consistent equations $Ax=y$ when $R(A)=n$
A_R^{-1}		$AG = I$	solving consistent equations $x'A=y'$ when $R(A)=m$
A^-		$AGA = A$	solving consistent equations
A_r^-		$AGA = A,\ GAG = G$	solving consistent equations
A_m^-	(i)	$AGA = A,\ (GA)^{\#} = GA$	minimum norm solution
	(ii)	$GA = P_{A^{\#}}$	
A_ℓ^-	(i)	$AGA = A,\ (AG)^{\#} = AG$	least squares solution
	(ii)	$AG = P_A$	
A^+	(i)	$AGA = A,\ GAG = G$ $(GA)^{\#} = GA,\ (AG)^{\#} = AG$	minimum norm least squares solution
	(ii)	$AG = P_A,\ GA = P_G$	

A matrix which is both a minimum norm g-inverse and a least squares g-inverse of A is denoted by $A_{\ell m}^-$. We some times consider two positive definite matrices M and N of orders $m \times m$ and $n \times n$ respectively and consider the specific inner products $(x,y) = x^*My\ \forall\ x,y \in E^m$ and $(x,y) = x^*Ny\ \forall\ x,y \in E^n$. In this case, minimum norm g-inverse, least squares g-inverse and minimum norm least squares g-inverse of A are denoted by $A^{-}_{m(N)}$, $A^{-}_{\ell(M)}$ and A^{+}_{MN} respectively. The detailed conditions which these g-inverses satisfy are given in table 4.

Table 4 : Basic types of inverses

Notation	Equivalent conditions	purpose
$A^{-}_{m(N)}$	(i) $(AG)^* = P_{A^*}$ (ii) $AGA = A$, $(GA)^*N=NGA$	minimum N-norm[x^*Nx] solution of consistent equation $Ax = y$
$A^{-}_{\ell(M)}$	(i) $AG = P_A$ (ii) $AGA = A$, $(AG)^*M=MAG$	M-least squares [$(Ax-y)^*M(Ax-y)$] solution of inconsistent equations $Ax = y$
A^{+}_{MN}	(i) $AG = P_A$, $GA = P_G$ (ii) $AGA = A$, $GAG = G$ $(GA)^*N = NGA$, $(AG)^*M = MAG$	minimum N-norm and M-least squares solution of $Ax = y$

C H A P T E R 1

CALCULUS OF VARIOUS TYPES OF g-INVERSES OF MATRICES

1.1 Introduction and Summary.

Over the past decade and a half several types of generalized inverses of singular and rectangular matrices have been developed depending either upon the uses or upon the properties which they are required to possess similar to those of the inverse of a nonsingular matrix. Along with the introduction of several new types of g-inverses, development of a calculus of these inverses has also been in progress. These studies have unfolded newer properties of these g-inverses hitherto unknown, established interconnections among the different types of g-inverses and probed into computational aspects of the g-inverses. In this chapter, we obtain several results on the calculus of g-inverses of matrices.

In section 1.2, we give a characterisation of Moore-Penrose inverse, the most well-known among all generalized inverses. Motivated by the beautiful result of Rao and Mitra (1970), '' if $G = A^-$ then G is A^-_{ℓ} if and only if $GG^{\#} = (A^{\#}A)^-$ '', we investigate in section 1.3 into the conditions $GG^{\#}$ should satisfy so that $G = A^-_{\ell m}$ or A^+

under the assumption that $G = A^-$. In this section we also obtain some results connecting g-inverses of A with those of powers of $A^{\#}A$ and $AA^{\#}$. In section 1.4 necessary and sufficient conditions on A are obtained so that every g-inverse of A is a minimum norm (least squares) g-inverse of A.

In section 1.5, we determine the class of all g-inverses with power property of a matrix A when $R(A) = R(A^2)$ and settle a conjecture of Mitra (1968) on g-inverses with specified manifolds.

We show that a matrix is uniquely determined by the class of its reflexive/least squares/minimum norm g-inverses, in section 1.6. Using this we provide an alternative proof of the result that a matrix is uniquely determined by the class of its g-inverses. (This result is due to Rao and Mitra, 1970).

In section 1.7, a theorem of Rao (1967) on projections is extended. An inner product is explicitly specified which converts virtually disjoint subspaces into pairwise orthogonal subspaces.

Section 1.8 is devoted to obtain a few results on nnd matrices. Here the class of all nnd g-inverses of a nnd matrix is determined. In section 1.9 a few matrix equations of interest are solved and an application of one of the results to the distribution of quadratic forms is suggested.

1.2 A characterisation of Moore-Penrose inverse.

Several characterisations of the celebrated Moore-Penrose inverse are well-known (Moore, 1920, 1935; Penrose, 1955; Householder, 1964; Rao, 1967; Mitra, 1968a - only to mention a few.). In this section another characterisation of Moore-Penrose inverse of A is given through a g-inverse of A and that of a power of $A^{\#}A$. In this section we consider the inner products $(x,y)_m = x^*My$ and $(x,y)_n = x^*Ny$ in E^m and E^n respectively. Accordingly, $A^{\#}$ is defined by the condition $(Ax, y)_m = (x, A^{\#}y)_n \; \forall \; x \in E^n$ and $y \in E^m$ where inner products are specified as above. We prove

Theorem 1 : Let A be a $m \times n$ matrix. Then a $n \times m$ matrix G is A^+ if and only if one of the following equivalent conditions holds.

(i) $G = A_r^-$ and $(GG^{\#})^p = [(A^{\#}A)^p]^-$

for some positive integer $p \geq 2$.

(ii) $G = A_r^-$ and $G^{\#}(GG^{\#})^p = [(A^{\#}A)^p\ A^{\#}]^-$ for some positive integer p.

Proof : The 'only if' part not only holds but is indeed too modest as it is well-known that if $G = A^+$, then G is clearly A_r^- and $(GG^{\#})^p = [(A^{\#}A)^p]^+$ and $G^{\#}(GG^{\#})^p = [(A^{\#}A)^p\ A^{\#}]^+$ for all positive integers p.

To prove the 'if' part, consider a general singular value decomposition of A (See Rao and Mitra, 1970),

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^* \qquad (1.2.1)$$

where U and V are matrices of order $m \times m$ and $n \times n$ respectively such that $U^*MU = I$ and $V^*N^{-1}V = I$ and D is a diagonal matrix of order $r \times r$ with diagonal elements as the positive square roots of the nonnull eigen values of $A^{\#}A$ where $r = R(A)$. Now,

$$G = A_r^- \iff G = N^{-1}V \begin{bmatrix} D^{-1} & J \\ F & FDJ \end{bmatrix} U^*M$$

$$\iff G^{\#} = U \begin{bmatrix} D^{-1} & F^* \\ J^* & J^*DF^* \end{bmatrix} V^* \qquad (1.2.2)$$

Define for $s = 1, 2, \ldots$

$$X_s = D^{-\frac{1}{2}+s} \cdot Y_s \, D^{\frac{5}{2}-s} \tag{1.2.3}$$

where $Y_s = JJ^*$ if s is odd and F^*F if s is even.

We need the following lemmas.

<u>Lemma 1</u>: Let G be as in (1.2.2). Then for every positive integer s,

$$(GG^{\#})^s = N^{-1}V \begin{bmatrix} D^{-2s} + D^{-2s+\frac{3}{2}}\{(I + X_{2s-1})\ldots(I+X_1)-I\}D^{-\frac{3}{2}} & \cdot \\ F\,D^{\frac{5}{2}-2s}\{(I + X_{2s-1})\ldots(I+X_1)\}D^{-\frac{3}{2}} & \cdot \end{bmatrix} V^*$$

and

$$G^{\#}(GG^{\#})^s = U \begin{bmatrix} D^{-2s-1} + D^{-2s+\frac{1}{2}}\{(I+X_{2s})\ldots(I+X_1)-I\}D^{-\frac{3}{2}} & \cdot \\ J^*D^{\frac{3}{2}-2s}\{(I+X_{2s})\ldots(I+X_1)\}D^{-\frac{3}{2}} & \cdot \end{bmatrix} V^*$$

where $\cdot$ represents an unspecified submatrix in the last $n-r$ columns.

Lemma 1 is established by induction on s.

<u>Lemma 2</u> : The matrices X_s, $s = 1,2, \ldots$ defined in (1.2.3) are semisimple with real nonnegative eigen values.

<u>Proof</u>: Observe that Y_s and hence $D Y_s D$ are hermitian and nnd for each s. Therefore for each s, there exists a unitary matrix L_s such that $L_s^* D Y_s D L_s = \Delta_s$ where Δ_s is a real diagonal matrix. Define $P_s = D^{-\frac{3}{2}+s} L_s$. It is easy to check that $P_s^{-1} X_s P_s = \Delta_s$. This completes the proof of lemma 2.

Coming back to the proof of the main theorem, now observe that

$$(A^{\#} A)^p = N^{-1} V \begin{bmatrix} D^{2p} & 0 \\ 0 & 0 \end{bmatrix} V^*$$

and $$(A^{\#} A)^p A^{\#} = N^{-1} V \begin{bmatrix} D^{2p+1} & 0 \\ 0 & 0 \end{bmatrix} U^* M$$

Hence in view of lemma 1

$(GG^{\#})^p = [(A^{\#} A)^p]^-$ and $G = A_r^- \Rightarrow$

$$D^{-2p+\frac{3}{2}} \{(I + X_{2p-1}) \ldots (I + X_1) - I\} D^{-\frac{3}{2}} = 0 \Rightarrow$$

$$(I + X_{2p-1}) \ldots (I + X_1) = I \Rightarrow \prod_{i=1}^{2p-1} | I + X_i | = 1 \Rightarrow$$

$X_i = 0$ for $i = 1, 2, \ldots, 2p-1$ in view of lemma 2.

Since $p \geq 2$, $X_1 = 0$ and $X_2 = 0$. Now $X_1 = 0 \Rightarrow J = 0$; and $X_2 = 0 \Rightarrow F = 0$. This completes the proof of 'if' part of (i). The proof of 'if' part of (ii) follows on similar lines as that of (i).

Thus theorem 1 is established.

<u>Corollary</u> : Let A be a $m \times n$ matrix. $G = A^+$ if and only if one of the following equivalent conditions holds.

(i) $G = A_r^-$ and $(G^{\#} G)^p = [(A A^{\#})^p]^-$ for some positive integer $p \geq 2$.

(ii) $G = A_r^-$ and $G(G^{\#} G)^p = [(A A^{\#})^p A]^-$ for some positive integer p.

<u>Proof</u>: The corollary follows trivially from theorem 1 once it is observed that (a) $G = A_r^- \iff G^{\#} = (A^{\#})_r^-$ and (b) $G = A^+ \iff G^{\#} = (A^{\#})^+$.

One can ask the following question. If in conditions (i) and (ii) of theorem 1, $G = A_r^-$ is replaced by $G = A^-$ (any g-inverse of A), <u>ceteris paribus</u>, will G be $A_{\ell m}^-$? The answer is in the negative. In fact G, then, may be neither A_ℓ^- nor A_m^-. This is exhibited in the following example. Take

$$M = N = I, A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \text{ and } G = \begin{bmatrix} 1 & 1 \\ 0 & \frac{3}{2} \end{bmatrix}.$$

A straightforward computation shows that $G = A^-$, $G^{\#}GG^{\#} = (A^{\#}AA^{\#})^-$ but $G \neq A^-_{\ell}$ and $G \neq A^-_m$.

Theorem 1 tells us the following. Let $G = A^-_r$. If further $(GG^{\#})^p = [(A^{\#}A)^p]^-$ for some positive integer $p \geq 2$ or $G^{\#}(GG^{\#})^p = [(A^{\#}A)^p A^{\#}]^-$ for some positive integer p, then indeed $(GG^{\#})^s = [(A^{\#}A)^s]^-$ and $G^{\#}(GG^{\#})^s = [(A^{\#}A)^s A^{\#}]^-$ for all positive integers s.

1.3 Connection between g-inverses of A and those of powers of $A^{\#}A$ and $AA^{\#}$.

The motivation for the results presented in this section is the following beautiful theorem of Rao and Mitra (1970). " Let $G = A^-$. Then $G = A^-_{\ell}$ if and only if $GG^{\#} = (A^{\#}A)^-$." We prove

Theorem 2 : Let $G = A^-$. Then

(a) $G = A^-_{\ell}$ if and only if $GG^{\#} = (A^{\#}A)^-$

(b) $G = A^-_{\ell r}$ if and only if $GG^{\#} = (A^{\#}A)^-_r$

(c) $G = A^-_{\ell m}$ if and only if $GG^{\#} = (A^{\#}A)^-_{\ell m}$

(d) $G = A^+$ if and only if $GG^{\#} = (A^{\#}A)^+$.

Proof of (a)* : 'Only if' part follows trivially once it is observed that $G = A^-_\ell \Rightarrow A^{\#}AG = A^{\#}$. To prove the 'if' part we proceed as follows. Let $G = A^-$. Then $GG^{\#} = (A^{\#}A)^-$ $\Rightarrow (A^{\#}AG - A^{\#})(A^{\#}AG - A^{\#})^{\#} = A^{\#}AGG^{\#}A^{\#}A - A^{\#}AGA - A^{\#}G^{\#}A^{\#}A + A^{\#}A = 0 \Rightarrow A^{\#} = A^{\#}AG \Rightarrow G = A^-_\ell$. This completes the proof of (a).

Proof of (b): Proof of (b) is complete in the light of (a) once it is observed that $R(G) = R(A) \Leftrightarrow R(GG^{\#}) = R(A^{\#}A)$.

Proof of (c): For the 'only if' part observe that

$$G = A^-_{\ell m} \Rightarrow A^{\#}AA^{\#}AGG^{\#} = A^{\#}AA^{\#}G^{\#} = A^{\#}A$$

$$\Rightarrow GG^{\#} = (A^{\#}A)^-_\ell \Rightarrow GG^{\#} = (A^{\#}A)^-_m.$$

This completes the proof of 'only if' part.

For the 'if' part, we first observe that $G = A^-$ and $GG^{\#} = (A^{\#}A)^- \Rightarrow G = A^-_\ell$ in view of (a). Further,

$$G = A^-_\ell \text{ and } GG^{\#} = (A^{\#}A)^-_\ell \Rightarrow A^{\#}AA^{\#}AGG^{\#} = A^{\#}A$$

$$= A^{\#}AA^{\#}G^{\#} \Rightarrow G = A^-_m.$$

This completes the proof of (c).

---

* As mentioned in the beginning of the section, theorem 2(a) is due to Rao and Mitra (1970). We reproduce their proof for completeness.

<u>Proof of (d)</u>: (d) follows trivially from (b) and (c). This completes the proof of theorem 2.

<u>Corollary</u>: Let A be a $m \times n$ matrix. Let the inner products in E^m and E^n be as defined in section 1.2. Let $G = A^-$. Then

(a) $G = A^-_{\ell(M)}$ if and only if $GM^{-1}G^* = (A^*MA)^-$.

(b) "$G = A^-_{\ell(M)}$ and $G = A^-_{m(N)}$" if and only if

$GM^{-1}G^* = (A^*MA)^-_{\ell(N^{-1})}$.

(c) $G = A^+_{MN}$ if and only if $GM^{-1}G^* = (A^*MA)^+_{N^{-1}N}$.

The corollary follows trivially from theorem 2. The proof of the following theorem follows in the same lines as that of theorem 2. We state

<u>Theorem 3</u>: Let $G = A^-$. Then

(a) $G = A^-_m$ if and only if $G^{\#}G = (AA^{\#})^-$

(b) $G = A^-_{\ell m}$ if and only if $G^{\#}G = (AA^{\#})^-_{\ell m}$

(c) $G = A^+$ if and only if $G^{\#}G = (AA^{\#})^+$.

We now prove

<u>Theorem 4</u>: (a) $G = A^-_{\ell m} \Rightarrow G^{\#}(GG^{\#})^p = [(A^{\#}A)^p A^{\#}]^-_{\ell m}$ for all positive integers p.

(b) $G^{\#}(GG^{\#})^p = [(A^{\#}A)^p A^{\#}]^-$ for some positive integer p and $G^{\#}G = (AA^{\#})^-_{\ell m} \Rightarrow G = A^-_{\ell m}$.

<u>Proof of (a)</u> : $G = A^-_{\ell m} \Rightarrow AA^{\#}AA^{\#}AA^{\#}G^{\#}GG^{\#} = AA^{\#}AA^{\#}G^{\#} = AA^{\#}A$. Hence the result is true for $p = 1$. Suppose the result is true for $p = 1,\ldots,p_o$. We shall show that then the result is true for $p = p_o + 1$.

$$G = A^-_{\ell m} \Rightarrow A(A^{\#}A)^{p_o+1} (A^{\#}A)^{p_o+1} A^{\#}(G^{\#}G)^{p_o+1} G^{\#}$$

$$= (AA^{\#})^2 \quad A(A^{\#}A)^{p_o} (A^{\#}A)^{p_o} A^{\#}(G^{\#}G)^{p_o} G^{\#}GG^{\#}$$

$$= (AA^{\#})^2 A(A^{\#}A)^{p_o} GG^{\#} = A(A^{\#}A)^2 (A^{\#}A)^{p_o-1}$$

$$= A(A^{\#}A)^{p_o+1}.$$

Proof of the Statement "$G = A^-_{\ell m} \Rightarrow G^{\#}(GG^{\#})^p = [(A^{\#}A)^p A^{\#}]^-_m$ follows on similar lines and the details are omitted.

Thus (a) is established.

<u>Proof of (b)</u>: Observe that $G^{\#}G = (AA^{\#})^-_{\ell m} \Rightarrow (G^{\#}G)^p = [(AA^{\#})^p]^-_{\ell m} \Rightarrow (AA^{\#})^p (G^{\#}G)^p = (G^{\#}G)^p (AA^{\#})^p \Rightarrow A^{\#}(AA^{\#})^p (G^{\#}G)^p = A^{\#}$.

Hence, $(G^{\#}G) = (AA^{\#})^-_{\ell m}$ together with $G^{\#}(GG^{\#})^p = [(A^{\#}A)^p A^{\#}]^- \Rightarrow (A^{\#}A)^p A^{\#} = (A^{\#}A)^p A^{\#}G^{\#}(GG^{\#})^p(A^{\#}A)^p A^{\#} = A^{\#}G^{\#}(A^{\#}A)^p A^{\#} \Rightarrow G = A^-$.

An appeal to theorem 3(b) now completes the proof of (b).

This completes the proof of theorem 4.

<u>Note</u> : The condition $G^{\#}G = (AA^{\#})^{-}_{\ell m}$ in theorem 4(b) can be replaced by the condition $GG^{\#} = (A^{\#}A)^{-}_{\ell m}$.

We conclude this section with the following remark. In theorem 4(b) the condition, "$G^{\#}G = (AA^{\#})^{-}_{\ell m}$" cannot be replaced by a weaker condition. In fact, $G^{\#}G = (AA^{\#})^{-}$ and $G^{\#}GG^{\#} = (A^{\#}AA^{\#})^{-}$ need not even imply that $G = A^{-}$. This is demonstrated in the following example. Take

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} \frac{1}{2} & \sqrt{3/20} \\ 0 & \sqrt{3/80} + \sqrt{3/16} \end{bmatrix}, \quad M = N = I.$$

1.4 When is every g-inverse a minimum norm (least squares) g-inverse ?

Rao (1967) stated that if the rank of a $m \times n$ matrix A is n, then every g-inverse of A is a minimum norm g-inverse of A. In this section we obtain a necessary and sufficient condition for every g-inverse of A to be a minimum norm (or least squares) g-inverse of A. We need

<u>Lemma 3</u> : Let A be a $m \times n$ matrix. Then there exists a $n \times m$ matrix U such that

$$\text{(a)}\quad AU = 0 \quad \text{and} \quad \text{(b)}\quad UA \neq 0$$

if and only if $1 \leq R(A) \leq n-1$.

<u>Proof</u> : Proof of 'only if' part is trivial. To prove the 'if' part we proceed as follows. Let $R(A) = r$ where $1 \leq r \leq n-1$. There exist two nonsingular matrices B and C such that

$$A = B \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} C .$$

Define a $n \times m$ matrix U as

$$U = C^{-1} \begin{bmatrix} 0_{r \times r} & 0 \\ E & F \end{bmatrix} B^{-1}$$

where E is nonnull and F is arbitrary.

Clearly, $AU = 0$ and $UA \neq 0$.

This completes the proof of lemma 3.

<u>Corollary</u> : Let A be a $m \times n$ matrix. There exists a $n \times m$ matrix U such that

$$\text{(a)}\quad AU \neq 0 \quad \text{and} \quad \text{(b)}\quad UA = 0$$

if and only if $1 \leq R(A) \leq m-1$.

The corollary trivially follows from lemma 3.

We now prove

<u>Theorem 5</u> : Let A be a $m \times n$ matrix. Every g-inverse of A is A_m^- if and only if $R(A)$ is either o or n.

<u>Proof</u> : Proof of 'if' part follows trivially once it is observed that if $R(A) = n$ then every g-inverse of A is a left inverse of A. To prove the 'only if' part we proceed as follows.

If every g-inverse of A is a minimum norm g-inverse of A, we have

$$(G + U - GAUAG)\ AA^{\#} = A^{\#} \qquad (1.4.1)$$

where G is a g-inverse of A and U is arbitrary.

From (1.4.1) it now follows that

$$(I - GA)\ UAA^{\#} = 0. \qquad (1.4.2)$$

Let if possible $1 \leq R(A) \leq n-1$. Then by lemma 3 there exists U_o such that $AU_o = 0$ but $U_o A \neq 0$. From (1.4.2), it now follows that $U_o\, AA^{\#} = GAU_o AA^{\#} = 0$ which in turn implies that $U_o A = 0$.

This is a contradiction since we chose $U_o \ni U_o A \neq 0$.

Hence R(A) is either n or o.

This completes the proof of theorem 5.

The following corollary is easy to prove.

Corollary : Let A be a $m \times n$ matrix. Then every g-inverse of A is a least squares g-inverse of A if and only if R(A) is either m or o.

Incidentally the main theorem of this section can also be proved using general singular value decomposition of a matrix. We omit the details.

1.5 A conjecture of Mitra on g-inverses with specified manifolds.

Let G be a g-inverse of a matrix A such that either $\mathcal{M}(G) = \mathcal{M}(A)$ or $\mathcal{M}(G^*) = \mathcal{M}(A^*)$. Mitra (1968b) proved that in such a case, for every positive integer p, $A^p G^p A^p = A^p$ and $G^p A^p G^p = G^p$. (Henceforth we shall find it convenient to call a g-inverse G of A with this property as a g-inverse of A with power property.) He then raised the following question: ''If G is a g-inverse of A with power property does it follow that either $\mathcal{M}(G) = \mathcal{M}(A)$ or $\mathcal{M}(G^*) = \mathcal{M}(A^*)$? ''

The answer in general is clearly in the negative. If $R(A) \neq R(A^2)$, then it is known (Mitra, 1968b) that there does not exist a g-inverse G of A such that either $\mathcal{M}(G) = \mathcal{M}(A)$ or $\mathcal{M}(G^*) = \mathcal{M}(A^*)$. However, the Scroggs-Odell pseudoinverse of A, say G, indeed exists for all matrices A and possesses the power property (See Scroggs and Odell, 1966).

We shall now give a counterexample even in the case where $R(A) = R(A^2)$. Before we do this, we determine the class of all g-inverses of A with power property where $R(A) = R(A^2)$. Before we proceed any further, we recall that if $R(A) = R(A^2)$, the Jordan representation of A can be written as $A = L \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} L^{-1}$ where C and L are nonsingular matrices. We prove

<u>Theorem 6</u> : Let $R(A) = R(A^2)$. Then G is a g-inverse of A with power property if and only if $G = L \begin{bmatrix} C^{-1} & J \\ F & FCJ \end{bmatrix} L^{-1}$ where F and J are arbitrary subject to the condition that $JF = 0$, C is nonsingular and $A = L \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} L^{-1}$

(As for example in Jordan representation of A.)

Proof :

Proof of 'if' part.

We need

Lemma 4 : Let G be as defined in Theorem 6. Then

$$G^p = L \begin{bmatrix} (C^{-1})^p & (C^{-1})^{p-1}J \\ F(C^{-1})^{p-1} & F(C^{-1})^{p-2}J \end{bmatrix} L^{-1}$$

for all positive integers p. Where $(C^{-1})^{-1}$ and $(C^{-1})^0$ are interpreted as C and I respectively.

Proof : Lemma 4 follows by induction on p.

In view of lemma 4, 'if' part of the theorem follows by straightforward computation.

Proof of 'only if' part.

Observe that any g-inverse G of A can be expressed as $G = L \begin{bmatrix} C^{-1} & J \\ F & H \end{bmatrix} L^{-1}$ where J, F and H are arbitrary. Now,

$$GAG = G \Rightarrow H = FCJ \text{ ; and } A^2G^2A^2 = A^2 \Rightarrow JF = 0.$$

This completes the proof of theorem 6.

The following corollary is easily established.

<u>Corollary</u> : Let A be a $n \times n$ matrix. If $R(A) = R(A^2) = n-1$ and $A = L \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} L^{-1}$, then every g-inverse G of A with power property can be written as $G = L \begin{bmatrix} C^{-1} & J \\ F & 0 \end{bmatrix} L^{-1}$ where either J or F or both are null.

From theorem 6, it clearly follows that the class of g-inverses G such that either $\mathcal{M}(G) = \mathcal{M}(A)$ or $\mathcal{M}(G^*) = \mathcal{M}(A^*)$ is only a subclass of the class of all g-inverses with power property in the case $R(A) = R(A^2)$. However if $R(A) = R(A^2) = n-1$, then these two classes are identical. Thus the answer to Mitra's query is in the affermative in the case where $R(A) = R(A^2) = n-1$.

We conclude this section with the following interesting remark. Let G be a g-inverse of A such that $\mathcal{M}(G) = \mathcal{M}(A)$. Then Mitra (1968b) proves that

$$\left.\begin{aligned} A^{r_1} G^{r_2} &= G^{r_2 - r_1} \\ \text{and} \quad G^{r_1} A^{r_2} &= A^{r_2 - r_1} \end{aligned}\right\} \text{ for all positive integers } r_1 \text{ and } r_2 \text{ such that } r_2 > r_1. \tag{1.5.1}$$

It is easy to see that if a g-inverse G of A satisfies the stronger property given in (1.5.1) then clearly $\mathcal{M}(G) = \mathcal{M}(A)$. A similar result can be stated for the case $\mathcal{M}(G^*) = \mathcal{M}(A^*)$.

1.6 A characterisation of a matrix by the class of its reflexive g-inverses.

Rao and Mitra (1970) proved that if $\{A^-\} = \{B^-\}$ then $A = B$. It is well-known that $A^+ = B^+ \Rightarrow A = B$. In this section we show that $\{A_r^-\} = \{B_r^-\} \Rightarrow A = B$; and $\{A_\ell^-\} = \{B_\ell^-\} \Rightarrow A = B$. Corresponding result for minimum norm g-inverses follows as a simple corollary. Using the results of this section an alternative proof of the result of Rao and Mitra is also presented. We prove

Theorem 7 : $A = B$ if and only if $\{A_r^-\} = \{B_r^-\}$.

Proof : 'Only if' part is trivial. To prove the 'if' part we proceed as follows. Observe that $\{A_r^-\} = \{B_r^-\} \Rightarrow R(A)=R(B)$. Let $A = CD$ and $B = EF$ be rank factorisations of A and B respectively. By lemma 2.5.2 of Rao and Mitra (1970), $G \in \{A_r^-\} \Rightarrow G = D_R^{-1} C_L^{-1}$ where D_R^{-1} is a right inverse of D and C_L^{-1} is a left inverse of C. Similarly $G \in \{B_r^-\} \Rightarrow G = F_R^{-1} E_L^{-1}$.

Now, $\{A_r^-\} = \{B_r^-\} \Rightarrow$

$$F\, D_R^{-1}\, C_L^{-1} E = I \tag{1.6.1}$$

and
$$D\, F_R^{-1}\, E_L^{-1} C = I \tag{1.6.2}$$

irrespective of the choices of right and left inverses of matrices involved.

Further (1.6.2) suggests that $F_R^{-1}\, E_L^{-1} C$ is one choice of D_R^{-1} substituting which in (1.6.1), we have

$$E_L^{-1} C\, C_L^{-1} E = I\,. \tag{1.6.3}$$

We note that for one choice of C_L^{-1}, $C\, C_L^{-1} = P_C$, the orthogonal projection operator onto $\mathcal{M}(C)$. Putting $E_L^{-1} = (E^*E)^{-1} E^*$ in (1.6.3), we have, after some simplification,

$$\begin{aligned} &E^*\, P_C E = E^*E \Rightarrow E = P_C E \\ &\Rightarrow E = CS \end{aligned} \tag{1.6.4}$$

where S is nonsingular.

We have, similarly,

$$D = TF \tag{1.6.5}$$

where T is nonsingular.

Substituting (1.6.4) and (1.6.5) in (1.6.2) we have

$$TS^{-1} = I \quad \text{or} \quad T = S.$$

Hence $A = CD = CTF = CSF = EF = B$.

This completes the proof of theorem 7.

Corollary (Rao and Mitra, 1970): $A = B$ if and only if $\{A^-\} = \{B^-\}$.

Proof : 'Only if' part is trivial. To prove the 'if' part, observe that $C \in \{A_r^-\} \Rightarrow C \in \{A^-\} \Rightarrow C \in \{B^-\}$. Hence $C \in \{A_r^-\}$ and $D \in \{B^-\} \Rightarrow R(D) \geq R(C) = R(A)$. In particular if $C = A_r^-$ and $D = B_r^-$, then $R(D) \geq R(C) = R(A)$. By a similar argument, it follows that $R(C) \geq R(D) = R(B)$. Hence $R(A) = R(B)$. Clearly, $R(A) = R(B)$ and $\{A^-\} = \{B^-\} \Rightarrow \{A_r^-\} = \{B_r^-\}$. An appeal to theorem 7 establishes the corollary.

Theorem 8 : $A = B$ if and only if $\{A_{\ell r}^-\} = \{B_{\ell r}^-\}$.

Proof : The 'only if' part is trivial. To prove the 'if' part observe that

$$\{A_{\ell r}^-\} = \{B_{\ell r}^-\} \Rightarrow R(A) = R(B).$$

Further $\{A^-_{\ell r}\} = \{B^-_{\ell r}\} \Rightarrow A^*A(B^*B)^-B^* = A^*$

$$\Rightarrow \mathcal{M}(A) = \mathcal{M}(B) \Rightarrow P_A = P_B$$

$$\Rightarrow A(B^*B)^-B^*B = P_A B = P_B B = B$$

which is thus invariant under the choice of a g-inverse of (B^*B). By result (iie) in section 1b.5 of Rao (1965) the invariance holds if and only if $A(B^*B)^-B^*B = A$. Hence $A = B$.

The following corollaries are easy to establish.

Corollary 1 : $A = B$ if and only if $\{A^-_\ell\} = \{B^-_\ell\}$

Corollary 2 : $A = B$ if and only if $\{A^-_m\} = \{B^-_m\}$

Corollary 3 : $A = B$ if and only if $\{A^-_{mr}\} = \{B^-_{mr}\}$

1.7 Extension of a theorem of Rao on projection.

''Let Λ be a positive definite matrix and let A_i be $m \times n_i$ matrix of rank r_i, $i = 1,2,\ldots,k$ such that $\Sigma R(A_i) = m$. Then the two statements

(a) $A_i^* \Lambda A_j = 0$ for all i, j such that $i \neq j$

and (b) $\Lambda^{-1} = \Sigma A_i (A_i^* \Lambda A_i)^- A_i^*$

are equivalent.'' This is a theorem in Rao (1967). Putting $\Lambda = I$ we get virtually a restatement of the celebrated Cochran's theorem which is of great importance in distribution of quadratic forms in normal variables.

In this section, we give an extension of the above theorem of Rao and also provide an inner product which converts virtually disjoint subspaces into orthogonal subspaces. In this section $\text{diag}(A_1, A_2, \ldots, A_k)$ represents the partitioned matrix i-th diagonal block of which is A_i, $i = 1, 2, \ldots, k$ all the other blocks being null matrices.

$$\left.\begin{array}{l}\text{Let } \Lambda \text{ be a } m \times n \text{ matrix. Let } A_i, B_i \text{ be} \\ \text{matrices of order } m \times p_i,\ n \times q_i \text{ respectively,} \\ i = 1, 2, \ldots, k. \text{ Write } A = (A_1 \vdots \ldots \vdots A_k) \text{ and} \\ B = (B_1 \vdots \ldots \vdots B_k).\end{array}\right\} \qquad (1.7.1)$$

Consider the following statements :

(a) $A_i^* \Lambda B_j = 0$ for all i and j such that $i \neq j$

(b) $G = \sum_i B_i (A_i^* \Lambda B_i)^- A_i^*$ is a g-inverse of Λ

where $(A_i^* \Lambda B_i)^-$ is any g-inverse of $A_i^* \Lambda B_i$

(c) $R(A^* \Lambda B) = R(\Lambda)$

and (d) $\sum_i R(A_i^* \Lambda) = \sum_i R(\Lambda B_i) = R(\Lambda)$

We prove

Theorem 9 : Consider the set up as in (1.7.1). Then (a) $\Rightarrow$ (b) if and only if (c) holds.

Proof:

Proof of 'if' part.

$$(a) \Rightarrow A^* \wedge B = \text{diag}((A_1^* \wedge B_1), \ldots, (A_k^* \wedge B_k))$$

$$\Rightarrow D = \text{diag}((A_1^* \wedge B_1)^-, \ldots, (A_k^* \wedge B_k)^-)$$

is a g-inverse of $(A^* \wedge B)$ whatever be the choices of $(A_i^* \wedge B_i)^-$ for $i = 1,2,\ldots,k$.

Hence, (a) and (c) in view of theorem 2.1 of Mitra (1968b) imply that $G = BDA^* = \sum_i B_i(A_i^* \wedge B_i)^- A_i^*$ is a g-inverse of $\wedge$ where $(A_i^* \wedge B_i)^-$ is any g-inverse of $A_i^* \wedge B_i$, for $i = 1,2,\ldots,k$.

This completes the proof of 'if' part.

Proof of 'only if' part :

Again in view of theorem 2.1 of Mitra (1968b),

$$B\,\text{diag}((A_1^* \wedge B_1)^-, \ldots, (A_k^* \wedge B_k)^-)A^* = \wedge^- \Rightarrow R(A^* \wedge B) = R(\wedge) \tag{1.7.2}$$

This completes the proof of theorem 9.

<u>Theorem 10</u>: Consider the set up as in (1.7.1). Then b) ⇒ (a) if and only if (d) holds.

<u>Proof</u> :

<u>Proof of 'if' part</u>.

Observe that (b) ⇒ $R(\Lambda) = tr\,(G\,\Lambda)$

$$= \sum_i tr(B_i(A_i^* \Lambda B_i)^- A_i^* \Lambda) = \sum_i tr((A_i^* \Lambda B_i)^- A_i^* \Lambda B_i) = \sum_i R(A_i^* \Lambda B_i). \qquad (1.7.3)$$

(1.7.3) and (d) together imply that

$$R(A_i^* \Lambda B_i) = R(A_i^* \Lambda) = R(\Lambda B_i) \text{ for } i = 1, 2, \ldots, k \qquad (1.7.4)$$

Let $D = diag\,((A_1^* \Lambda B_1)^-, \ldots, (A_k^* \Lambda B_k)^-)$.

Observe next that

$$(b) \Rightarrow \Lambda BDA^* \Lambda = \Lambda \Rightarrow A^* \Lambda BDA^* \Lambda B = A^* \Lambda B$$

$$\Rightarrow D = (A^* \Lambda B)^-. \qquad (1.7.5)$$

We need

<u>Lemma 5</u> : Consider the set up as in (1.7.1). If (d) holds and (b) is true for some choice of g-inverses of $(A_i^* \Lambda B_i)$, $i = 1, 2, \ldots, k$, it is true for every choice.

Proof: Consider the equality

$$\Lambda = \Lambda G \Lambda = \sum_i \Lambda B_i (A_i^* \Lambda B_i)^- A_i^* \Lambda \qquad (1.7.6)$$

and observe that for each i the matrix $\Lambda B_i (A_i^* \Lambda B_i)^- A_i^* \Lambda$ is invariant under the choice of g-inverses of $(A_i^* \Lambda B_i)$ in view of (1.7.4). This completes the proof of lemma 5.

Henceforth, without loss of generality (in view of lemma 5) choose and fix g-inverses occuring in D so that $R((A_i^* \Lambda B_i)^-) = R(A_i^* \Lambda B_i)$. Note that in view of (1.7.2) and (1.7.3),

$$R(D) = \sum_i R(A_i^* \Lambda B_i)^- = \sum_i R(A_i^* \Lambda B_i) = R(\Lambda) = R(A^* \Lambda B) \qquad (1.7.7)$$

Now, (1.7.5) and (1.7.7) in view of theorem 2a of Mitra (1968a) imply that $DA^* \Lambda BD = D$.

$$\text{Therefore, } (A_i^* \Lambda B_i)^- (A_i^* \Lambda B_j)(A_j^* \Lambda B_j)^- = 0 \quad \text{whenever } i \neq j \qquad (1.7.8)$$

(1.7.8) => whenever $i \neq j$

$$(A_i^* \Lambda B_i)(A_i^* \Lambda B_i)^- (A_i^* \Lambda B_j)(A_j^* \Lambda B_j)^- (A_j^* \Lambda B_j) = 0$$

=> whenever $i \neq j$, $A_i^* \Lambda B_j = 0$ in view of (2.3) using corollary 1a.3 of Mitra (1968a).

This completes the proof of 'if' part.

<u>Proof of 'only if' part.</u>

Assume that both (a) and (b) hold. Then,

$$A_i^* \Lambda = A_i^* \Lambda G \Lambda = A_i^* \Lambda \left(\sum_i B_i (A_i^* \Lambda B_i)^- A_i^*\right) \Lambda =$$

$$A_i^* \Lambda B_i (A_i^* \Lambda B_i)^- A_i^* \Lambda \Rightarrow B_i (A_i^* \Lambda B_i)^-$$ is a g-inverse of

$A_i^* \Lambda$.

Hence, $G \Lambda = \sum_i B_i (A_i^* \Lambda B_i)^- A_i^* \Lambda = \sum_i (A_i^* \Lambda)^- A_i^* \Lambda$.

This implies that $R(\Lambda) = tr(G \Lambda) = \sum_i tr((A_i^* \Lambda)^- A_i^* \Lambda) =$

$\sum_i R(A_i^* \Lambda)$.

Similarly we prove that $R(\Lambda) = \sum_i R(\Lambda B_i)$.

This completes the proof of theorem 10.

Now a few remarks are in order. If Λ is nonsingular and $A_i = B_i$, then (c) is equivalent to $R(A) = m$ and (d) is equivalent to $\sum_i R(A_i) = m$. When Λ is positive definite (d) is implied by (b) for

$$\Lambda^{-1} = \sum_i A_i (A_i^* \Lambda A_i)^- A_i^* \Rightarrow I = \sum_i A_i (A_i^* \Lambda A_i)^- A_i^* \Lambda$$

$$= \sum_i A_i A_i^- \Rightarrow m = tr\, I = \sum_i tr(A_i A_i^-) = \sum_i R(A_i).$$

Hence, Rao's theorem stated in the beginning of the section can be restated thus :

Theorem 11 : Let $R(A) = m$ and Λ be positive definite. Then (a) $A_i^* \Lambda B_j = 0$ for all i, j such that $i \neq j \Longleftrightarrow$ (b) $I = \sum_i A_i(A_i^* \Lambda A_i)^- A_i^* \Lambda$.

If $E^m = \mathcal{M}(A_1) \oplus \mathcal{M}(A_2) \oplus \ldots \oplus \mathcal{M}(A_k)$ represents a decomposition of E^m into mutually orthogonal subspaces with inner product defined as $(x,y)=x^* \Lambda y$ and if $P_i = A_i(A_i^* \Lambda A_i)^- A_i^* \Lambda$ denotes the orthogonal projection operator projecting arbitrary vectors onto $\mathcal{M}(A_i)$ (See lemma 3 of Mitra and Rao, 1968), then theorem 11(b) asserts the well-known result that such projection operators P_i add to I.

Now let

$$E^m = \mathcal{M}(A_1) \oplus \mathcal{M}(A_2) \oplus \ldots \oplus \mathcal{M}(A_k) \qquad (1.7.9)$$

represent an arbitrary decomposition of E^m. We shall obtain a positive definite matrix Λ such that for $x, y \in E^m$ if inner product is defined as $x^* \Lambda y$, then $\mathcal{M}(A_1), \ldots, \mathcal{M}(A_k)$ become pairwise orthogonal subspaces.

Now we prove

<u>Theorem 12</u>: Let $\Lambda = (\sum_{i=1}^{k} A_i A_i^*)^{-1}$ where $A_1, \ldots, A_k$ are as in (1.7.9). Then (i) Λ is positive definite and (ii) $A_i^* \Lambda A_j = 0$ whenever $i \neq j$.

<u>Proof</u> : Clearly Λ is positive definite since $R(A_1 : \ldots : A_k) = m$. Observe that Λ is a g-inverse of $A_i A_i^*$ for $i = 1,2,\ldots,k$. (See complement 5 after section 1b in Rao, 1965). Hence $A_i^*(\sum_{j=1}^{k} A_j A_j^*)^{-1} A_i = A_i^*(A_i A_i^*)^- A_i$ for $i = 1,2,\ldots,k$. Observe that for each i $A_i^*(A_i A_i^*)^- A_i$ is invariant under the choice of g-inverse of $A_i A_i^*$ and further it is idempotent. Thus $A_i^*(A_i A_i^*)^- A_i$ is a g-inverse of itself. Therefore for one choice of $(A_i^* (\sum_{j=1}^{k} A_j A_j^*)^{-1} A_i)^-$ we have

$$A_i(A_i^* (\sum_{j=1}^{k} A_j A_j^*)^{-1} A_i)^- A_i^* = A_i A_i^*(A_i A_i^*)^- A_i A_i^*$$

$$= A_i A_i^* .$$

Thus, $\sum_{i=1}^{k} A_i(A_i^* \Lambda A_i)^- A_i^* = \sum_{i=1}^{k} A_i A_i^* = \Lambda^{-1}$. Now an appeal to theorem 11 shows that $A_i^* \Lambda A_j = 0$ whenever $i \neq j$.

This completes the proof of theorem 12.

1.8 g-inverses of nonnegative definite matrices.

In this section, we obtain the most general form of a nnd g-inverse of a nnd matrix. We also prove a few more results of related interest.

Theorem 13 : A hermitian matrix A has a nnd g-inverse if and only if A is nnd.

Proof : Theorem 13 is easy to establish and we omit the proof.

We now prove

Theorem 14 : (Most general form of nnd g-inverse): Let $A=MM^*$ be a nnd matrix of order $m \times m$ where M is a $m \times r$ matrix of rank r. Then G is a nnd g-inverse of A if and only if $G = K^*K$ where $K = LM^- + U(I-MM^-)$

where n is an arbitrary integer $\geq r$,

L is an arbitrary $n \times r$ matrix such that $L^*L = I_r$,

U is an arbitrary $n \times m$ matrix,

and M^- is any g-inverse of M.

Proof : To prove the 'if' part observe that since M is a $m \times r$ matrix of rank r, every g-inverse of M is a left inverse of M. Straightforward verification establishes that K^*K is a g-inverse of M. Further K^*K is nnd. This completes the proof of 'if' part.

To prove the 'only if' part we proceed as follows. Let $G = K^*K$ be a g-inverse of $A = MM^*$ where K is a $n \times m$ matrix of rank n. Clearly, $n \geq r$. Now $MM^*K^*KMM^* = MM^* \Rightarrow M^*K^*KM = I_r$. Thus $K = KMM^- + K(I-MM^-)$ where KM is a matrix such that $M^*K^*KM = I_r$. Hence putting $L = KM$ and $U = K$, it is easily verified that the 'only if' part holds. This completes the proof of theorem 14.

We now go on to prove the last theorem of this section.

<u>Theorem 15</u> : Let A be a nnd matrix of order $m \times m$ and X be a $s \times m$ matrix of rank s. Then XAX^* is idempotent if and only if X can be expressed as $X = YC$ where

$$Y = L(N^*N)^{-1} N^* + U(I-N(N^*N)^{-1} N^*),$$

C is an arbitrary matrix of order $p \times m$ and of rank p where $p \geq s$,

N is a $p \times t$ matrix satisfying the equation $CAC^* = NN^*$ where $t = R(N) = R(CAC^*)$,

L is an arbitrary matrix of order $s \times t$ such that $L^*L = I_t$,

and U is an arbitrary matrix of order $s \times p$ such that $R(Y) = s$.

Proof : 'If' part follows by straightforward verification.

For the 'only if' part choose $C = X$, $L = N$ and $U = I$ and observe that with such a choice X can indeed be expressed in the required form.

1.9 Solution of matrix equations.

In this section, we solve the matrix equations $XAXA = XA$, $XAX = 0$ and $XAXAX = XAX$ and make a few comments on some solutions of matrix equations which are already known.

Lemma 6 : The most general form of an idempotent matrix H of order $n \times n$ is given by $H = C^- C$ where C is an arbitrary $m \times n$ matrix, m being arbitrary.

Proof of lemma 6 is omitted as it is straightforward.

We prove

Theorem 16 : Let A be a $m \times n$ matrix. Then XA is idempotent if and only if X is of the form

$$X = (CA)^{-}C + E(A^{\perp})^{*}$$

where $p, q \geq m-R(A)$ are arbitrary positive integers and C, E and $(A^{\perp})$ are arbitrary matrices of order $p \times m$, $n \times q$ and $m \times q$ respectively such that $A^{*}A^{\perp} = 0$.

Proof : $((CA)^{-}C + E(A^{\perp})^{*})A = (CA)^{-}CA$ is clearly idempotent. Now if XA is idempotent, XA is a g-inverse of itself. Choose $p = n$, $C = X$, $(CA)^{-} = XA$, we have $(CA)^{-}C = XAX$. It is now clear that $(X-(CA)^{-}C)A = (X-XAX)A = 0$. This completes the proof of theorem 16.

Theorem 17 : Let A be a hermitian matrix of order $m \times m$. Then for a hermitian matrix X, XA is idempotent if and only if X is of the form

$$C^{*}(CAC^{*})^{-}_{r}\, C + A^{\perp} D(A^{\perp})^{*}$$

where $p, q \geq m-R(A)$ are arbitrary, C, $A^{\perp}$ are arbitrary matrices of order $p \times m$ and $m \times q$ respectively such that $A^{*}A^{\perp} = 0$, D is an arbitrary diagonal matrix of order $q \times q$ and $(CAC^{*})^{-}_{r} = G(CAC^{*})G^{*}$ where G is a g-inverse of CAC^{*}.

Proof : 'If' part follows by straightforward verification. To prove the 'only if' part observe that if XA is idempotent AXA is a hermitian g-inverse of XAX. Choosing $C = X$, $G = AXA$, $(CAC^*)^-_r = AXAXAXAXA = AXA$ we have, $C^*(CAC^*)^-_r C = XAXAX = XAX$. Now the rest of the proof follows on the same lines as in theorem 16.

This completes the proof of theorem 17.

Theorem 18: For a matrix A of order $m \times n$, $XAX = 0$ if and only if X is a $n \times m$ matrix of the form $X = YC$ where p is an arbitrary positive integer, C is an arbitrary matrix of order $p \times m$ and Y is an arbitrary solution of $CAY = 0$.

Proof : 'If' part follows trivially. To prove the 'only if' part let X_o be a solution of the equation $XAX = 0$ and let $R(X_o) = p$. Let $X_o = YC$ be a rank factorisation of X_o. Now $YCAYC = 0 \Rightarrow CAY = 0$. This completes the proof of theorem 18.

Theorem 19 : For matrices A and W of orders $m \times n$ and $q \times m$ respectively, $XAX = 0$ and $WAX = 0$ if and only if X is a $n \times m$ matrix of the form $X = YC$ where p is an arbitrary positive integer, C

is an arbitrary $p \times m$ matrix and Y is an arbitrary solution of the equation $\begin{bmatrix} C \\ W \end{bmatrix} AY = 0$.

Proof of theorem 19 is similar to that of theorem 18 and we omit.

Now we prove

<u>Theorem 20</u> : For a matrix A of order $m \times n$, $XAXAX = XAX$ if and only if $X = Z + W$ where W is a solution of the equation $WAW = W$ and Z satisfies the equations $ZAZ = 0$ and $WAZ = 0$.

<u>Proof</u> :

<u>Proof of 'if' part.</u>

Let Z and W be as given in the hypothesis. Then $(Z+W)A(Z+W) = ZAZ + WAZ + ZAW + WAW = ZAW + W$ and

$(Z+W)A(Z+W)A(Z+W) = (ZAW + W)A(Z+W)$

$= ZAWAZ + ZAWAW + WAZ + WAW$

$= ZAW + W.$

Hence $X = Z+W$ satisfies the equation $XAXAX = XAX$.

<u>Proof of 'only if' part.</u>

Let X be a solution of the equation $XAXAX = XAX$. Observe that $W = XAX$ satisfies the equation $WAW = W$ and

$$(X-XAX)A(X-XAX) = XAX - XAXAX - XAXAX + XAXAXAX = 0$$

Also $WA(X-XAX) = XAXAX - XAXAXAX = 0.$

This completes the proof of theorem 20.

Mitra (1968a) showed that the most general solution of the equation $WAW = W$ is given by

$$W = Q(PAQ)_r^- P$$

where P and Q are arbitrary.

Using theorems 19 and 20 we therefore observe that the most general solution of the equation $XAXAX = XAX$ is given by

$$X = Z + W$$

where $W = Q(PBQ)_r^- P$

and $Z = YC$

C, P and Q being arbitrary and Y being an arbitrary solution of the equation $\binom{C}{W}BY = 0$. It is well-known ((Ogasawara and Takahashi 1951; Khatri, 1963; Rao, 1965) that if a vector valued random variable x follows a m-variate normal distribution with null mean vector and dispersion matrix Σ, the quadratic form $x'Ax$ has a chi-square distribution if and only if

$$\Sigma A \Sigma A \Sigma = \Sigma A \Sigma.$$

Given a symmetric matrix A, theorems 14 and 19 are useful in determining the class of all n.n.d. matrices Σ for which the chi-square distribution holds for $x'Ax$. Mitra (1968a) solved the dual problem where given Σ, he determined the class of all matrices A satisfying the equation $\Sigma A \Sigma A \Sigma = \Sigma A \Sigma$.

The following theorem is due to Morris and Odell (1968). We state the theorem as given in their paper and make a few comments on this useful theorem.

<u>Theorem 21</u> : (Morris and Odell). Let n be a positive integer and A_i be a $p \times q$ matrix and B_i be a $p \times r$ matrix for $i = 1,2,\ldots,n$. Define $C_1 = A_1$, $D_1 = B_1$, $E_1 = A_1^+ B_1$ and $F_1 = I - A_1^+ A_1$. Furthermore define $C_k = A_k F_{k-1}$, $D_k = B_k - A_k E_{k-1}$, $E_k = E_{k-1} + F_{k-1} C_k^+ D_k$ and $F_k = F_{k-1}(I - C_k^+ C_k)$ for $k = 2,3,\ldots,n$. Then $A_i X = B_i$, for $i = 1,2,\ldots,n$ have a common solution if and only if $C_i C_i^+ D_i = D_i$ for $i = 1,2,\ldots,n$. In case there exists a

solution, the general common solution is $X = E_n + F_n Z$ where Z is arbitrary.

We wish to point out that in the above theorem Moore-Penrose inverses are of no importance. In fact, the theorem as stated above holds even if the Moore-Penrose inverse is replaced by any g-inverce(in the sense that is understood in this thesis) wherever it occurs. In the latter case also proof follows exactly on the same line as that of theorem 21. Further, A_i and B_i can be taken as $p_i \times q$ and $p_i \times r$ matrices respectively, $i = 1, 2, \ldots, n$. This is to say that A_i and A_j need not have the same number of rows when $i \neq j$.

Making these modifications theorem 21 can now be stated as

Theorem 22: Let n be a positive integer and let A_i and B_i be $p_i \times q$ and $p_i \times r$ matrices for $i = 1, 2, \ldots, n$. Define $C_1 = A_1$, $D_1 = B_1$, $E_1 = A_1^- B_1$, $F_1 = (I - A_1^- A_1)$. Furthermore, define $C_k = A_k F_{k-1}$, $D_k = B_k - A_k E_{k-1}$, $E_k = E_{k-1} - F_{k-1} C_k^- D_k$ and $F_k = F_{k-1} (I - C_k^- C_k)$ for $k = 2, \ldots, n$. Then $A_i X = B_i$ for $i = 1, 2, \ldots, n$ have a common solution if and only if $C_i C_i^- D_i = D_i$ for $i = 1, 2, \ldots, n$. In case a common solution exists, the general common solution is given by $X = E_n + F_n Z$ where Z is arbitrary.

Proof of theorem 22 follows in the same lines as those of theorem 21 which is due to Morris and Odell and we omit the proof.

Notice that the conditions in theorem 22 are easier to verify in general as one can compute any g-inverse and check the conditions in theorem 22 where as in theorem 21 each time one has to compute Moore-Penrose inverse. As mentioned in the general introduction, this is one situation where any g-inverse serves as good a purpose as M-P inverse.

CHAPTER 2

g-INVERSES OF PARTITIONED MATRICES AND APPLICATIONS

2.1 Introduction and Summary.

Greville (1960) gave an interesting formula connecting A^+ with $(A : a)^+$ where a is a column vector. The same paper described a use of this result in least squares polynomial curve fitting where the degree of the polynomial has to be determined by fitting polynomials of successive degrees in stages until a satisfactory fit is obtained. In the absence of a table of orthogonal polynomials Greville's formula should prove to be highly expedient. Cline (1964) gave formulae to obtain A^+ from $(A : B)^+$. These formulae though mathematically elegant do not seem to be easily computable when B itself contains more than one column as several other Moore-Penrose inverses have to be computed in the process (though of lower order.)

In this chapter, we develop formulae similar to those of Greville and Cline connecting g-inverses of A with the corresponding g-inverses of $(A : a)$ where a is a column vector. We consider A^-, A_r^-, A_ℓ^-, A_m^- and A^+ for this purpose. In sections 2 and 3 two parallel sets of formulae are developed, one for computing $(A : a)^-$ from A^- and the other for computing A^- from $(A : a)^-$. In sections 2 and 3 we consider only the Euclidean norms. In section 4, the results of sections 2 and 3 are generalised where we consider more general inner products.

The results of sections 2, 3 and 4 are useful in least squares computations. First let us consider a linear model $Y = X\beta + \varepsilon$ where $D(\varepsilon) = \sigma^2 I$. The results of section 2 are useful when one looks for revised estimates of parameters when either an additional uncorrelated observation with unit variance is considered or an extra parameter is added to the linear model. The results of section 3 are useful when one wishes to compute revised estimates of the remaining parameters if a superfluous parameter is dropped out from the original linear model. These are also useful in the dual problem when an outlying observation wrongly considered has to be removed from the analysis and one has to work out the consequent corrections in the least squares estimates. Now consider a linear model $Y = X\beta + \varepsilon$ where $D(\varepsilon) = \mathbf{A}\sigma^2$ where $\mathbf{A}$ is a known positive definite matrix. The results of section 4 are useful in making the recalculation of least squares estimates for data or model changes in this case. These applications are discussed in detail in section 6.

In section 5 we discuss g-inverses of a sum of two matrices. In section 7 application of the results of the first few sections to least squares recalculations is illustrated by a few examples.

In section 8 we extend a theorem of Rohde (1965) on g-inverses of partitioned matrices. The object of this section is only to show that Rohde's results are true for a much wider class of matrices.

2.2 Derivation of g-inverses of (A : a) from those of A.

In this section, we obtain several types of g-inverses of (A : a) from the corresponding types of g-inverses of A in two mutually exclusive and collectively exhaustive cases, namely, $a \in \mathcal{M}(A)$ and $a \notin \mathcal{M}(A)$.

Let G be a g-inverse of a $m \times n$ matrix A and $a \in E^m$. Then it is well-known (Rao, 1967) that $a \in \mathcal{M}(A) \iff AGa = a$. Thus given a g-inverse of A, one can check, using the above mentioned result, whether a vector a belongs to $\mathcal{M}(A)$ or not. Now we shall consider

Case 1 : $a \notin \mathcal{M}(A)$

$$\left.\begin{array}{l} \text{Let } A \text{ be a } m \times n \text{ matrix and } G \text{ be a} \\ \text{g-inverse of } A. \text{ Let } a \in E^m \text{ but } a \notin \mathcal{M}(A). \\ \text{Define } d = Ga,\ c = (I-AG)^* (I-AG)a, \\ b = \dfrac{c}{c^*a} \text{ and } X^* = (G^* - bd^* : b). \end{array}\right\} \quad (2.2.1)$$

We prove

Lemma 1 : Consider the set up in (2.2.1). c is a nonnull vector and c^*a is real and positive.

Proof : $c^*a = a^*(I-AG)^* (I-AG)a$ is clearly real and nonnegative. Further $c = 0 \iff c^*a = 0 \iff (I - AG)a = 0 \iff a \in \mathcal{M}(A)$. Hence under the set up (2.2.1) $c^*a \neq 0$ and $c \neq 0$. This completes the proof of lemma 1.

Note : In view of lemma 1, definition of b in (2.2.1) is valid.

We now prove

<u>Theorem 1</u> : Consider the set up in (2.2.1). Then

(i) $X = (A : a)^-$

(ii) $X = (A : a)_r^-$ if $G = A_r^-$

(iii) $X = (A : a)_m^-$ if and only if $G = A_m^-$

(iv) $X = (A : a)_\ell^-$ if $G = A_\ell^-$

and (v) $X = (A : a)^+$ if $G = A^+$.

<u>Proof</u>:

<u>Proof of (i) and (ii)</u> : Observe that by construction $b^*A = 0^*$ and $b^*a = 1$. Now a straightforward verification completes the proof of (i). Proof of (ii) is similar to that of (i).

<u>Proof of (iii)</u> : It suffices to show that $X(A : a)$ is hermitian. Now,

$$X(A : a) = \begin{bmatrix} GA - d\, b^*A & Ga - d\, b^*a \\ b^*A & b^*a \end{bmatrix}$$

Recall that $b^*A = 0^*$. Further $Ga = d\, b^*a$ since $b^*a = 1$. Hence $X = (A : a)_m^- \iff (GA)^* = GA \iff G = A_m^-$.

Proof of (iv) and (v) : $(A : a)X = AG - Adb^* + ab^* =$
$AG + (I - AG)\, a\, a^*(I-AG)^*\,(I-AG)/c^*a = AG + (I-AG)aa^*(I-AG)^*/c^*a$,
since $G = A_\ell^-$.

Observe that AG is hermitian and so is
$(I-AG)aa^*(I-AG)^*/c^*a$ in view of lemma 1. Thus $(A : a)X$ is hermitian. This completes the proof of (iv). (v) is similarly established.

Remark 1 : (v) of theorem 1 is due to Greville (1960).

Remark 2 : In the case of (iv) and (v) of theorem 1, c is reduced to the form $(I-AG)a$ since $G = A_\ell^-$.

We now prove

Theorem 2 : Consider the set up in (2.2.1). Then
(i) $X = (A : a)_r^- \Rightarrow G = A_r^-$ if and only if $G\,a \in \mathcal{M}(GA)$
(ii) $X = (A : a)_\ell^- \Rightarrow G = A_\ell^-$ if and only if $(AG)^*a = (AG)^*AGa$
and (iii) $X = (A : a)^+ \Rightarrow G = A^+$ if and only if $G\,a \in \mathcal{M}(GA)$ and $(AG)^*a = (AG)^*A\,G\,a$.

Proof :

Proof of (i) : $X = (A : a)_r^-$ and $G\,a \in \mathcal{M}(GA) \Rightarrow$
$GAG - GAGab^* + Gab^* = G \Rightarrow GAG = G$. This completes the proof of 'if' part. Now, $X = (A : a)_r^-$ and $G = A_r^-$
$\Rightarrow GAGab^* = Gab^* \iff GAGa = Ga \iff Ga \in \mathcal{M}(GA)$.

This completes the proof of 'only if' part.

Proof of (ii) : $(AG)^*a = (AG)^* AGa \Rightarrow (AG)^*(I-AG)a = 0$ $\Rightarrow (I-AG)a = (I-AG)^* (I-AG)a \Rightarrow (I-AG)aa^*(I-AG)^*(I-AG)/c^*a$ $= (I-AG)\, aa^*(I-AG)^*/c^*a$ is hermitian. Hence, $X = (A : a)^-_\ell$ and $(AG)^*a = (AG)^* AGa \Rightarrow AG$ is hermitian. This completes the proof of 'if' part. Now $G = A^-_\ell \Rightarrow (AG)^* AGa = AGAGa =$ AGa. This completes the proof of 'only if' part.

(iii) follows from theorem 1 and (i) and (ii) of theorem 2.

Now we shall consider

Case 2 : $a \in \mathcal{M}(A)$.

We prove

Theorem 3 : Let G be a g-inverse of A and $a \in \mathcal{M}(A)$. Let $d = Ga$. Then the following hold.

(i) $\begin{bmatrix} G - db^* \\ b^* \end{bmatrix} = (A : a)^-$ where b is an arbitrary vector.

(ii) $\begin{bmatrix} G - db^* \\ b^* \end{bmatrix} = (A : a)^-_r$ if and only if $G = A^-_r$ and $b \in \mathcal{M}(G^*)$.

(iii) Let b be an arbitrary vector. Then

$$\begin{bmatrix} G-db^* \\ b^* \end{bmatrix} = (A : a)_{\ell}^{-} \text{ if and only if } G = A_{\ell}^{-} .$$

(iv) Let $b = G^*Ga/(1 + a^*G^*Ga)$. Then $\begin{bmatrix} G-db^* \\ b^* \end{bmatrix} = (A : a)_{m}^{-}$ if and only if $G = A_{m}^{-}$.

(v) Let b be as defined in (iv). Then

$$\begin{bmatrix} G-db^* \\ b^* \end{bmatrix} = (A : a)^{+} \text{ if and only if } G = A^{+}.$$

Proof:

Proof of (i): Observe that $a \in \mathcal{M}(A) \Rightarrow AGa = a$.

Hence $(A : a) \begin{bmatrix} G-db^* \\ b^* \end{bmatrix} = AG - Adb^* + ab^* = AG$ where b is arbitrary. Further $AG(A : a) = (A : a)$. Notice that this holds for every vector b. This completes the proof of (i).

Proof of (ii) and (iii) :

$$\begin{bmatrix} G-db^* \\ b^* \end{bmatrix} AG = \begin{bmatrix} GAG - db^*AG \\ b^*AG \end{bmatrix} \qquad (2.2.2)$$

Now $G = A_{r}^{-}$ and $b \in \mathcal{M}(G^*) \Leftrightarrow GAG = G$ and $b^*AG = b^*$

$\Leftrightarrow \begin{bmatrix} G - db^* \\ b^* \end{bmatrix} = (A : a)_{r}^{-}$ in view of (2.2.2). This completes the proof of (ii).

(iii) is a trivial consequence of the fact that

$$A\bar{G} = (A \vdots a) \begin{bmatrix} G - db^* \\ b^* \end{bmatrix} \quad \text{where } b \text{ is arbitrary.}$$

Proof of (iv) :

Proof of 'if' part : $a \in \mathcal{M}(A)$ and $G = A_m^- \Rightarrow GA$ is hermitian and $db^*A = \dfrac{G\,a\,a^*G^*G\,A}{1 + a^*G^*Ga} = \dfrac{G\,a\,a^*\,G^*}{1 + a^*G^*Ga} \Rightarrow$ $GA - db^*A$ is hermitian. Further, $a \in \mathcal{M}(A)$ and $G = A_m^- \Rightarrow$

$$b^*A = \frac{a^*\,G^*}{1 + a^*G^*Ga} = (Ga - db^*a)^*. \text{ Finally,}$$

$$b^*a = \frac{a^*G^*G\,a}{1 + a^*G^*Ga}$$ is a real number. Hence it follows that

$$\begin{bmatrix} G - d\,b^* \\ b^* \end{bmatrix} (A \vdots a)$$ is hermitian. This completes the proof of 'if' part.

Proof of 'only if' part : Observe that

$$\left.\begin{array}{l} \begin{bmatrix} G - d\,b^* \\ b^* \end{bmatrix} (A \vdots a) \text{ is hermitian } \Rightarrow GA - db^*A \\ \text{is hermitian and } (b^*A)^* = G\,a - d\,b^*a. \end{array}\right\} \quad (2.2.3)$$

$$\left.\begin{aligned}&\text{Now, } (b^*A)^* = Ga - db^*a \Rightarrow db^*A = G\,a\,a^*G^* + d\,a^*\,b\,d^* \\ &\Rightarrow d\,b^*A \text{ is hermitian.}\end{aligned}\right\} \quad (2.2.4)$$

From (2.2.3) and (2.2.4) it follows that

$$\begin{bmatrix} G - db^* \\ b^* \end{bmatrix} (A : a) \text{ is hermitian} \Rightarrow GA \text{ is hermitian.}$$

This completes the proof of (iv).

Proof of (v) follows from that of (i) - (iv).

This completes the proof of theorem 3.

<u>Remark 3</u> : 'If' part of (v) of theorem 3 is due to Greville (1960).

2.3 <u>Derivation of g-inverses of A from those of (A : a).</u>

In this section we obtain various types of g-inverses of A from the corresponding g-inverses of (A : a). Let $(G^* : b)^*$ be a g-inverse of $(A : a)$. We consider the problem in three mutually exclusive and collectively exhaustive cases namely, $a \notin \mathcal{M}(A)$, $a \in \mathcal{M}(A)$ and $b^*a \neq 1$, and $a \in \mathcal{M}(A)$, $b^*A \neq 0^*$ and $b^*a = 1$. (That these are indeed mutually exclusive and collectively exhaustive is established in lemma 2 given below.)

We prove

<u>Lemma 2</u> : Let $(G^* : b)^*$ be a g-inverse of $(A : a)$. Then $a \notin \mathcal{M}(A)$ if and only if $b^*A = 0^*$ and $b^*a = 1$.

<u>Proof</u>:

<u>Proof of 'if' part</u> : $(G^* : b)^* = (A : a)^- \Rightarrow$

$$AGA + a\,b^*\,A = A \tag{2.3.1}$$

and $$AGa + a\,b^*\,a = a \tag{2.3.2}$$

Now,

$b^*A = 0^*$ and $(2.3.1) \Rightarrow G = A^-$.

$b^*a = 1$ and $(2.3.2) \Rightarrow AGa = 0$ and $a \neq 0$.

From the above arguments it follows that

$(G^* : b)^* = (A : a)^-$, $b^*A = 0^*$ and $b^*a = 1 \Rightarrow a \notin \mathcal{M}(A)$

(for, $a \in \mathcal{M}(A) \Rightarrow A\,G\,a = a = 0$ which is a contradiction to the fact that $b^*a = 1$.)

This completes the proof of 'if' part.

<u>Proof of 'only if' part</u> :

$(2.3.2)$ and $a \notin \mathcal{M}(A) \Rightarrow AGa = a(1 - b^*a)$. and $a \notin \mathcal{M}($

$\Rightarrow 1 - b^*a = o \Rightarrow b^*a = 1$.

Now, $y \in \mathcal{M}(A) \Rightarrow (A : a)(G^* : b)^* y = y = A G y + a b^* y$

$\Rightarrow a b^* y \in \mathcal{M}(A)$ since $A G y$ and $y \in \mathcal{M}(A)$.

Further, $a \notin \mathcal{M}(A)$ and $a b^* y \in \mathcal{M}(A) \Rightarrow b^* y = o$.

From the above arguments it clearly follows that

$a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)^- \Rightarrow b^* A = 0^*$.

This completes the proof of lemma 2.

<u>Remark 4</u> : $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)^-$ $\Rightarrow A G a = 0$.

<u>Note</u> : Given a g-inverse $(G^* : b)^*$ of $(A : a)$, one can use lemma 2 to determine whether $a \in \mathcal{M}(A)$ or not.

We shall consider

<u>Case 1 : $a \notin \mathcal{M}(A)$</u>.

We prove

<u>Theorem 4</u> : Let $(G^* : b)^* = (A : a)^-$ and $a \notin \mathcal{M}(A)$. Then the following hold.

(i) $G = A^-$

(ii) Let $(G^* : b)^* = (A : a)_r^-$. Then $G = A_r^-$ if and only if $G a = 0$.

(iii) '$G = A_m^-$ and $Ga = 0$' if and only if $(G^* : b)^* = (A : a)_m^-$.

(iv) Let $(G^* : b)^* = (A : a)_\ell^-$. Then $G = A_\ell^-$ if and only if $a \varepsilon \mathfrak{G}(A)$.

(v) Let $(G^* : b)^* = (A : a)^+$. Then $G = A^+$ if and only if $a \varepsilon \mathfrak{G}(A)$.

Proof :

(i) is a trivial consequence of lemma 2.

Proof of (ii) : $G a = 0$, $(G^* : b)^* = (A : a)_r^- \Rightarrow$ $G A G + G a b^* := G = G A G$. This completes the proof of 'if' part. Again, $G = A_r^-$ and $(G^* : b)^* = (A : a)_r^- \Rightarrow$ $G a b^* = 0 \Rightarrow G a = G a b^* a = 0$. This completes the proof of 'only if' part.

Proof of (iii) : $(G^* : b)^* = (A : a)_m^- \Rightarrow GA$ is hermitian and $Ga = (b^*A)^* = 0$. This completes the proof of 'if' part. Again, $G = A_m^-$ and $G a = 0 \Rightarrow GA$ is hermitian and $0 = G a = (b^*A)^*$. Further $b^*a = 1 = (b^*a)^*$. Thus $G = A_m^-$ and $G a = 0 \Rightarrow (G^* : b)^* = (A : a)_m^-$. This completes the proof of 'only if' part.

Proof of (iv) :

Proof of 'if' part : From lemma 2 it follows that $a \notin \mathcal{M}(A) \Rightarrow b^*a = 1$ and $b^*A = 0^*$. Further $a \in \mathfrak{S}(A)$ $\Rightarrow A^*a = 0$. Now, $(G^* : b)^* = (A : a)^-_\ell \Rightarrow A\,G + a\,b^*$ is hermitian and G is a g-inverse of $A \Rightarrow A\,G = (AG + ab^*)AG$ $= (G^*A^* + b\,a^*)AG = G^*A^*\,AG \Rightarrow AG$ is hermitian. This completes the proof of 'if' part.

Proof of 'only if' part : $G = A^-_\ell \Rightarrow AG$ is the orthogonal projection operator which projects vectors onto $\mathcal{M}(A)$. Further from remark 4 it follows that $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)^- \Rightarrow A\,G\,a = 0$. From the above arguments it is clear that under the hypothesis $a \in \mathfrak{S}(A)$. This completes the proof of 'only if' part.

(v) follows from (i) - (iv).

This completes the proof of theorem 4.

Now a few remarks are in order.

Remark 5 : Let $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)^-$. Then $G = A^-_r \Rightarrow G\,a = 0$ and $(G^* : b)^* = (A : a)^-_r$.

Remark 6 : Consider the same set up as in remark 5. Then

(i) $G = A_\ell^- \Rightarrow a \in \mathcal{C}(A)$ and $(G^* : b)^* = (A : a)_\ell^-$

and (ii) $G = A^+ \Rightarrow a \in \mathcal{C}(A)$ and $(G^* : b)^* = (A : a)^+$.

To obtain A_r^-, A_ℓ^- and A^+ from $(A : a)_r^-$, $(A : a)_\ell^-$ and $(A : a)^+$ respectively one can use theorem 4 when $a \in \mathcal{C}(A)$. In the next two theorems we obtain these g-inverses for A from the corresponding ones of $(A : a)$ where a is an arbitrary vector not belonging to $\mathcal{M}(A)$.

We prove

<u>Theorem 5</u> : Let $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)_\ell^-$.

Then $GG^*\left(I + \dfrac{A^* a b^* G^*}{1 - b^* G^*A^*a}\right)A^*$ is one choice for A_ℓ^-.

<u>Proof</u> : From theorem 2(a) of Chapter 1 it follows that

$$(G^* : b)^* = (A : a)_\ell^- \Rightarrow \begin{bmatrix} GG^* & Gb \\ b^*G^* & b^*b \end{bmatrix} = \begin{bmatrix} A^*A & A^*a \\ a^*A & a^*a \end{bmatrix}^- \quad (2.$$

Further $a \notin \mathcal{M}(A) \Rightarrow (a^*A : a^*a)^* \notin \mathcal{M}[(A^*A : A^*a)^*]$

(2.

Hence (2.3.3) and (2.3.4) in view of theorem 4(i) and the fact that $(B^*)^- = (B^-)^*$ for any matrix B imply that

$(G^* : b)^* = (A : a)_\ell^-$ and $a \notin \mathcal{M}(A) \Rightarrow (GG^* : Gb)^* = (A^*A : A^*a)^-$ (2.3.5)

Observe that $A^*a \in \mathcal{M}(A^*A)$. Now, $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)_\ell^- \Rightarrow AG + ab^*$ is hermitian and $b^*a = 1 = b^*(AG+ab^*)a$ $\Rightarrow 1 = b^* G^* A^*a + b^*b\ a^*a$. Hence under the hypothesis, $b^*G^*A^*a = 1 \Rightarrow b^*ba^*a = 0 \Rightarrow$ either a or b is null vector $\Rightarrow$ $b^*a = 0$. This is a contradiction since under the hypothesis $b^*a = 1$.

Hence under the hypothesis $b^*G^*A^*a \neq 1$ (2.3.6)

We need lemma 3 which we state below and prove.

Lemma 3 : Let $(G^* : b)^* = (A : a)^-$ and $b^*a \neq 1$.

Then $G\left(I + \dfrac{ab^*}{1 - b^*a}\right)$ is a g-inverse of A.

Proof of the lemma : Under the hypothesis, it follows from (2.3.2) that $AGa = a(1 - b^*a)$. Further since $b^*a \neq 1$, it follows that $a = AGa/(1-b^*a)$. It now follows from (2.3.1) that $G\left(I + \dfrac{ab^*}{1-b^*a}\right) = A^-$.

This completes the proof of the lemma.

An appeal to lemma 3 leads in view of (2.3.5) and (2.3.6) to the result that under the hypothesis

$$GG^*\left(I + \frac{A^* a\, b^*\, G^*}{1 - b^* G^* A^* a}\right) = (A^* A)^-$$

which in turn implies that

$$GG^*\left(I + \frac{A^*\, a\, b^*\, G^*}{1 - b^* G^* A^* a}\right) \quad A^* = A_\ell^- \; . \, .$$

This completes the proof of theorem 5.

Remark 7 : The g-inverse of A constructed in theorem 5 is in fact $A_{\ell r}^-$.

We now prove

Theorem 6 : Let $a \notin \mathcal{M}(A)$ and $(G^* : b)^* = (A : a)^-$.

Then

(i) $G\left(I - \frac{b\, b^*}{b^* b}\right) = A^-$

(ii) $G\left(I - \frac{b\, b^*}{b^* b}\right) = A_r^-$ if $(G^* : b)^* = (A : a)_r^-$.

(iii) $G\left(I - \frac{b\, b^*}{b^*\, b}\right) = A_\ell^-$ if $(G^* : b)^* = (A : a)_\ell^-$.

(iv) $G\left(I - \frac{b\, b^*}{b^*\, b}\right) = A_m^-$ if $(G^* : b)^* = (A : a)_m^-$

and (v) $G\left(I - \frac{b\, b^*}{b^*\, b}\right) = A^+$ if $(G^* : b)^* = (A : a)^+$

Proof : (i) follows trivially.

Proof of (ii) : Under the hypothesis, $GAG + Gab^* = G$ and hence $G(I - \frac{b\,b^*}{b^*\,b})\,AG\,(I - \frac{b\,b^*}{b^*\,b}) = GAG(I - \frac{b\,b^*}{b^*\,b})$ $= (G - Gab^*)\,(I - \frac{b\,b^*}{b^*\,b}) = G(I - \frac{b\,b^*}{b^*\,b})$. This completes the proof of (ii).

Proof of (iii) : Under the hypothesis, $AG + ab^* = G^*A^* + b\,a^*$. To show that $AG\,(I - \frac{b\,b^*}{b^*\,b})$ is hermitian it suffices to show that $AG - G^*A^* - \frac{A\,G\,b\,b^*}{b^*\,b} + \frac{b\,b^*G^*A^*}{b^*\,b} = 0$.

Now, $AG - G^*A^* - \frac{A\,G\,b\,b^*}{b^*b} + \frac{b\,b^*\,G^*A^*}{b^*b} =$

$$ba^* - ab^* - \frac{(G^*A^* + ba^* - ab^*)\,bb^*}{b^*b} + \frac{bb^*(AG + ab^* - ba^*)}{b^*b} = 0.$$

This completes the proof of (iii).

Proof of (iv) : $G(I - \frac{bb^*}{b^*b})A = GA$ is indeed hermitian since $(G^* : b)^*\,(A : a)$ is hermitian. This completes the proof of (iv).

(v) follows from (i) - (iv). This completes the proof of theorem 6.

Remark 8 : (v) of theorem 6 is due to Cline (1964).

Remark 9 : Consider the same set up as in theorem 6. Then

(i) $G(I - \frac{b\,b^*}{b^*\,b}) = A_r^- \Rightarrow (G^* : b)^* = (A : a)_r^-$ if $a = \frac{(I - AG)b}{b^*b}$.

(ii) $G(I - \frac{b\,b^*}{b^*\,b}) = A_\ell^- \Rightarrow (G^* : b)^* = (A : a)_\ell^-$ if and only if $a = \frac{(I - AG)b}{b^*\,b}$.

(iii) $G(I - \frac{b\,b^*}{b^*b}) = A_m^-$ and $Ga = 0 \Longleftrightarrow$ $(G^* : b)^* = (A : a)_m^-$.

(iv) $G(I + \frac{b\,b^*}{b^*\,b}) = A^+ \Rightarrow (G^* : b)^* = (A : a)^+$ if $G\,a = 0$ and $a = \frac{(I - AG)b}{b^*b}$.

Remark 10 : Consider the same set up as in theorem 6. Then

(i) $G(I - ab^*) = A^-$

(ii) $G(I - ab^*) = A_r^-$ if and only if $(G^* : b)^* = (A : a)_r^-$

We shall now consider

Case 2 : $a \in \mathcal{M}(A)$ and $b^*a \neq 1$.

We prove

Theorem 7 : Let $(G^* : b)^* = (A : a)^-$ and $b^*a \neq 1$. Then the following hold :

(i) $G(I + \frac{ab^*}{1 - b^*a}) = A^-$

(ii) $G(I + \frac{ab^*}{1 - b^*a}) = A_r^-$ and $b^* = \frac{b^*AG}{1 - b^*a}$ if and only if $(G^* : b)^* = (A : a)_r^-$.

(iii) $G(I + \frac{ab^*}{1 - b^*a}) = A_m^-$, $(Ga)^* = b^*A$ and b^*a is real if and only if $(G^* : b)^* = (A : a)_m^-$.

(iv) $G(I + \frac{ab^*}{1 - b^*a}) = A_\ell^-$ if and only if $(G^* : b)^* = (A : a)_\ell^-$.

(v) $G(I + \frac{ab^*}{1 - b^*a}) = A^+$, $b^* = \frac{b^*AG}{1 - b^*a}$, $(Ga)^* = b^*A$ and b^*a is real if and only if $(G^* : b)^* = (A : a)^+$.

Proof :

Proof of (i) : Observe that $b^*a \neq 1 \Rightarrow a \in \mathcal{M}(A)$ and that (i) is the same as lemma 3.

<u>Proof of (ii)</u> : Observe that

$$(G^* : b)^* = (A : a)_r^- \Rightarrow \begin{cases} a = A\,G\,a/(1 - b^*a) & (2.3.7) \\ G = G\,A\,G + G\,a\,b^* & (2.3.8) \\ b^* = b^*A\,G + b^*a\,b^* & (2.3.9) \end{cases}$$

From (2.3.9) it clearly follows that

$$(B^* : b)^* = (A : a)_r^- \Rightarrow b^* = b^*A\,G/(1 - b^*a).$$

Now assume that $(G^* : b)^* = (A : a)_r^-$. Then, we have,

$$G(I + \frac{ab^*}{1-b^*a})\;AG(I + \frac{ab^*}{1-b^*a}) = G\,A\,G + \frac{Gab^*AG}{1-b^*a} + \frac{GAGab^*}{1-b^*a}$$

$$+ \frac{Gab^*AGab^*}{(1-b^*a)^2} = G(I + \frac{ab^*}{1-b^*a}).$$ This completes the proof of 'if' part. To prove the 'only if' part we proceed as follows. $b^* = b^*A\,G/(1-b^*a) \Rightarrow b^* = b^*A\,G + b^*a\,b^*$. Further, $(G^*:b)^* =$

$(A : a)^-$, $G(I + \frac{ab^*}{1 - b^*a}) = A_r^-$ and $b^* = \frac{b^*\,A\,G}{1 - b^*a}$ imply that

$$G = GAG + \frac{Gab^*AG}{1 - b^*a} + \frac{GAGab^*}{1 - b^*a} + \frac{Gab^*AGab^*}{(1 - b^*a)^2} - \frac{Gab^*}{1-b^*a} =$$

$$GAG + Gab^* + Gab^* + \frac{Gab^*ab^*}{1 - b^*a} - \frac{Gab^*}{1-b^*a} = GAG + Gab^*.$$ This completes the proof of 'only if' part.

Proof of (iii) : To prove the 'if' part we proceed as follows. $(G^* : b)^* = (A : a)^-_m \Rightarrow Ga = (b^*A)^*$ and b^*a is real. Now $(G^* : b)^* = (A : a)^-_m \Rightarrow GA + \frac{G\,a\,b^*A}{1 - b^*a}$ $= GA + \frac{(b^*A)^*\,b^*A}{1 - b^*a}$ is hermitian since GA is hermitian and b^*a is real. This completes the proof of 'if' part.

Now, assume that $(G^* : b)^* = (A : a)^-$, $G(I + \frac{ab^*}{1-b^*a}) = A^-_m$, $(G\,a)^* = b^*A$ and b^*a is real. To show that $(G^* : b)^* = (A : a)^-_m$ it suffices now to show that GA is hermitian. Under the hypothesis of the 'only if' part,

$GA + \frac{G\,ab^*A}{1 - b^*a}$ is hermitian and $\frac{Gab^*A}{1 - b^*a}$ is also hermitian. Hence under the hypothesis, GA is hermitian. This completes the proof of 'only if' part.

Proof of (iv) : Observe that $(G^* : b)^* = (A : a)^- \Rightarrow$ $a = \frac{AGa}{1 - b^*a} \Rightarrow AG + \frac{AGab^*}{1 - b^*a} = AG + ab^*$. Hence $G(I + \frac{ab^*}{1 - b^*a})$ $= A^-_\ell$ if and only if $(G^* : b)^* = (A : a)^-_\ell$. This completes the proof of (iv).

(v) follows from (i) - (iv).

This completes the proof of theorem 7.

<u>Remark 11</u> : 'If' part of (v) in theorem 7 is due to Cline (1964).

We shall finally consider

<u>Case 3 : $b^*a = 1$ and $b^*A \neq 0^*$</u> .

Let $(G^* : b)^*$ be a g-inverse of $(A : a)$ such that $b^*a = 1$ and $b^*A \neq 0^*$. Let the j-th column of A and the j-th row of G be denoted by a_j and g_j^* respectively. Since $b^*A \neq 0$, let its j-th coordinate $b^*a_j \neq o$. Define $c = a + a_j$ and let E be the matrix obtained from G by replacing its j-th row g_j^* by $g_j^* - b^*$. It is simple to check that $(E^* : b)^*$ is a g-inverse of $(A : c)$. Moreover if $(G^* : b)^*$ is a reflexive or least squares g-inverse of $(A : a)$ then so is $(E^* : b)^*$ of $(A : c)$. Observe that $b^*c = b^*a + b^*a_j \neq 1$ so that the methods of the previous case can be applied to obtain a g-inverse of A from $(E^* : b)$. Thus we arrive at the following

<u>Theorem 8</u> : Let $(G^* : b)^* = (A : a)^-$ such that $b^*a = 1$ and $b^*A \neq 0^*$. Let $b^*a_j \neq o$ (a_j is the j-th column of A). Let E be the matrix obtained from G by replacing its j-th row by 'its j-th row - b^*' keeping the other rows unaltered. Denote

$$X = E\left(I - \frac{(a + a_j)\, b^*}{b^* a_j}\right).$$

Then

(i) $X = A^-$

(ii) $X = A_r^-$ if $(G^* : b)^* = (A : a)_r^-$

and (iii) $X = A_\ell^-$ if $(G^* : b)^* = (A : a)_\ell^-$.

It is interesting to record that the two types of g-inverses left out in theorem 8, namely minimum norm and Moore-Penrose inverses never occur in the case under consideration. This we show in

<u>Theorem 9</u> : If $(G^* : b)^* = (A : a)_m^-$ then $b^* a = 1 \iff b^* A = 0^*$.

<u>Proof</u> : Observe that under the hypothesis $A^* b = Ga$. Now it is easy to check that $b^* a = 1 \iff AGa = 0 \iff AA^* b = 0 \iff A^* b = 0$. This completes the proof of theorem 9.

2.4 <u>Extension of the results in sections 2.2 and 2.3 to more general inner products.</u>

In this section we obtain minimum norm and least squares g-inverses of $(A : a)$ from those of A and vice versa considering more general inner products. Let M and N be positive definite matrices of orders $m \times m$ and $n \times n$ respectively. Let $N_1 = \begin{bmatrix} N & \theta \\ \theta^* & \delta \end{bmatrix}$ be a positive definite matrix of order $(n+1) \times (n+1)$ where N is as specified above. Define $(x,y)_m = x^*M\,y$ for $x,y \in E^m$, $(x,y)_n = x^*Ny$ for $x,y \in E^n$ and $(x,y) = x^*N_1\,y$ for $x,y \in E^{(n+1)}$. We use these notations throughout this section without explicit mention. We recall that $A^-_{\ell(M)}$, $A^-_{m(N)}$ denote respectively a M-least squares g-inverse and a minimum N-norm g-inverse of A where A is a $m \times n$ matrix. Throughout this section we assume that A is a $m \times n$ matrix and a, a $m \times 1$ vector. We only state the relevant results here without proof as they follow on similar lines to the corresponding ones (corresponding to Euclidean inner products) in the previous two sections.

<u>Theorem 10</u> : Let $G = A^-_{\ell(M)}$ and $a \notin \mathcal{M}(A)$. Define $d = Ga$, $c = (I - AG)a$, $b = c/c^*M\,a$ and $X = (G^* - Mbd^* : Mb)^*$. Then $X = (A : a)^-_{\ell(M)}$.

Theorem 11 : (a) Let $G = A^{-}_{m(N)}$ and $a \not\in \mathcal{M}(A)$. Define $d = Ga$, $c = (I - AG)^* (I - AG)a$, $b = c/c^*a$ and $X^* = (G^* - bd^* \vdots b)^*$. Then $X = (A \vdots a)^{-}_{m(N_1)}$ if and only if $\theta \in \mathcal{M}(A^*)$. (b) Let $G = A^{+}_{MN}$ and $a \not\in \mathcal{M}(A)$. Let d, b and X be as defined in theorem 10. Then $X = (A \vdots a)^{+}_{MN_1}$ if and only if $\theta \in \mathcal{M}(A^*)$.

Theorem 12 : (a) Let $G = A^{-}$ and $a \in \mathcal{M}(A)$. Define $d = Ga$. Then $(G^* - bd^* \vdots b)^* = (A \vdots a)^{-}_{\ell(M)}$ if and only if $G = A^{-}_{\ell(M)}$ where b is an arbitrary vector.

(b) Let $G = A^{-}$ and $a \in \mathcal{M}(A)$. Let $\theta = 0$ and $\delta = 1$. Define $d = Ga$ and $b = G^*N\,G\,a/(1 + a^*\,G^*\,N\,G\,a)$. Then $(G^* - bd^* \vdots b)^* = (A \vdots a)^{-}_{m(N_1)}$ if and only if $G = A^{-}_{m(N)}$.

(c) Consider the same set up as in (b). Then $X = (A \vdots a)^{+}_{MN_1}$ if and only if $G = A^{+}_{MN}$.

Theorem 13 : (a) Let $(G^* \vdots b)^* = (A \vdots a)^{-}_{\ell(M)}$ and $a \not\in \mathcal{M}(A)$. Then $G = A^{-}_{\ell(M)}$ if and only if $a \in \mathcal{C}(A)$.

(b) Let $(G^* \vdots b)^* = (A \vdots a)^{-}_{m(N_1)}$ and $a \not\in \mathcal{M}(A)$. Then $G = A^{-}_{m(N)}$.

(c) Let $(G^* : b)^* = (A : a)^+_{MN_1}$ and $a \not\in \mathcal{M}(A)$.

Then $G = A^+_{MN}$ if and only if $Ga = 0$ and $a^* MA = 0$.

(d) Let $(G^* : Mb)^* = (A : a)^-_{\ell(M)}$ and $a \not\in \mathcal{M}(A)$.

Then $G(I - \frac{bb^*M}{b^*Mb})$ is $A^-_{\ell(M)}$.

<u>Theorem 14</u> : (a) Let $(G^* : Mb)^* = (A : a)^-_{\ell(M)}$ and $b^*M\,a \neq 1$. Then $G(I + \frac{ab^*M}{1-b^*Ma})$ is $A^-_{\ell(M)}$.

(b) Let $(G^* : Mb)^* = (A : a)^-_{m(N_1)}$ and $b^*M\,a \neq 1$.

Then $G(I + \frac{ab^*M}{b^*Ma}) = A^-_{m(N)}$.

<u>Remark 12</u> : Let $(G^* : Mb)^* = (A : a)^-$. Then $a \not\in \mathcal{M}(A)$ if and only if $b^*Ma = 1$ and $b^*MA = 0^*$.

Thus, the case that is left out is the case where $b^*MA \neq 0^*$ and $b^*Ma = 1$. This is analogous to the corresponding case 3 in the previous section and the situation can be tackled in an exactly similar way and we omit the details.

2.5 g-inverse of a sum of two matrices.

Let A, B, C and D be matrices of orders $m \times n$, $m \times r$, $n \times r$ and $r \times r$ respectively. Further let D be nonsingular. Let G be a g-inverse of A. Denote $W = C^*GB + D^{-1}$ and assume that W is non-singular. Let $X = G - G\,B\,W^{-1}C^*G$. (2.5.1)

We prove

<u>Theorem 15</u> : (i) $X = (A+BDC^*)^-$ if either $\mathcal{M}(B) \subset \mathcal{M}(A)$ or $\mathcal{M}(C) \subset \mathcal{M}(A^*)$ or both.

(ii) $X = (A + BDC^*)_r^-$ if $G = A_r^-$ and either $\mathcal{M}(B) \subset \mathcal{M}(A)$ or $\mathcal{M}(C) \subset \mathcal{M}(A^*)$ or both.

(iii) $X = (A + BDC^*)_m^-$ if $\mathcal{M}(C) \subset \mathcal{M}(A^*)$ and $G = A_m^-$.

(iv) $X = (A+BDC^*)_\ell^-$ if $\mathcal{M}(B) \subset \mathcal{M}(A)$ and $G = A_\ell^-$.

(v) $X = (A+BDC^*)^+$ if $\mathcal{M}(B) \subset \mathcal{M}(A)$, $\mathcal{M}(C) \subset \mathcal{M}(A^*)$ and $G = A^+$.

<u>Proof</u> :

<u>Proof of (i)</u> : $G = A^-$ and $\mathcal{M}(B) \subset \mathcal{M}(A) \Rightarrow (A+BDC^*)X = AG + BDC^*G - B(C^*GB+D^{-1})^{-1}\,C^*G - BDC^*GB(C^*GB+D^{-1})^{-1}C^*G \doteq AG \Rightarrow (A+BDC^*)\,X\,(A+BDC^*) = AG(A+BDC^*) = A + BDC^*$ (Since $\mathcal{M}(B) \subset \mathcal{M}(A)$)

The proof is similar in the case when $\mathcal{M}(C) \subset \mathcal{M}(A^*)$. This completes the proof of (i).

Proof of (ii) - (v) follows from proof of (i) and the definitions of the g-inverses involved.

This completes the proof of theorem 15.

We now give several interesting corollaries of theorem 15.

Corollary 1 : Let $G = A^-$ and $B^*GB + D^{-1}$ be non-singular. Then $G - GB(B^*GB + D^{-1})^{-1}B^*G$ is a g-inverse of $A + BDB^*$ if either $\mathcal{M}(B) \subset \mathcal{M}(A)$ or $\mathcal{M}(B^*) \subset \mathcal{M}(A^*)$ or both.

Corollary 2 : Let A be a n.n.d. matrix and D be positive definite. Further let $\mathcal{M}(B) \subset \mathcal{M}(A)$ and G be a g-inverse of A. Then $B^*GB + D^{-1}$ is positive definite and $G - GB(B^*GB + D^{-1})^{-1}B^*G$ is a g-inverse of $(A + BDB^*)$.

Proof of Corollary 2 : Note that under the hypothesis B^*GB is invariant under the choices of g-inverses of A. Further, since A is n.n.d. there exists a n.n.d. g-inverse of A(See theorem 13 of Chapter 1). Hence B^*GB is n.n.d. and $D^{-1} + B^*GB$ is p.d. The rest of the proof follows from Corollary 1.

This completes the proof of Corollary 2.

Corollary 3 : (Rayner and Pringle, 1970). Consider the same set up as in corollary 2 with $D = I$. Then $G - GB(I + B^*GB)^{-1} B^*G$ is a g-inverse of $A + BB^*$.

Corollary 4 : (Rao and Mitra, 1970). Let $G = A^-$, a be a vector and d, a nonzero scalar. Assume that $\frac{1}{d} + a^*Ga \neq 0$. Then $\dfrac{G\,a\,a^*G}{\frac{1}{d} + a^*Ga}$ is a g-inverse of $A + d\,a\,a^*$ if either $a \varepsilon \mathcal{M}(A)$ or $a \varepsilon \mathcal{M}(A^*)$ or both.

Corollary 5 : (Rao, 1965). Let A be nonsingular and $D^{-1} + B^*A^{-1}B$ be also nonsingular. Then $A^{-1} + A^{-1}B\,(B^*A^{-1}B + D^{-1})^{-1}\,B^*A^{-1}$ is the inverse of $A + BDB^*$.

Remark 13 : To obtain g-inverses of A from those of $A + BDC^*$ one can use theorem 15 as $A = A+BDC^* + B(-D)C^*$.

2.6 Application of the results obtained in sections 2.2-2.4 to re-calculation of least squares estimates for data or model changes.

Consider a linear model $E(y)=X\beta$ and $D(y)=\sigma^2 I$ where y is a $n \times 1$ random vector of observations, X is a $n \times m$ matrix of known coefficients, β is a $m \times 1$ vector of parameters, and E(y) and D(y) denote the expectation and dispersion matrix of y respectively. (2.6.1)

$\hat{\beta}$ is said to be a least squares estimate of β if $(y - X\beta)'(y - X\beta)$ is minimum when $\beta = \hat{\beta}$. Clearly $\hat{\beta} = X_{\ell}^{-} y$ where X_{ℓ}^{-} is any least squares g-inverse of X. Further if $R(X) = r$, then $\hat{\sigma}^2 = \frac{R_o^2}{n - r}$ where $R_o^2 = y'y - y' X\hat{\beta}$.

Suppose X_{ℓ}^{-} is computed and $\hat{\beta}$ and $\hat{\sigma}^2$ are obtained from the formulae quoted above. We now consider two types of problems.

<u>Problem I</u> : An additional observation y_{n+1}, the expected value and variance of which are $\sum_{i=1}^{m} x_{n+1,i}\beta_i$ (where $x_{n+1,i}$ are known constants) and σ^2 respectively is given. Further y_{n+1} is uncorrelated with y_i, $i = 1,2,\ldots,n$. In the light of this new observation the estimates $\hat{\beta}$ and $\hat{\sigma}^2$ are to be revised.

To solve this problem, we proceed as follows. Denote $x' = (x_{n+1,1},\ldots,x_{n+1,m})$. Here the problem essentially is to compute $\begin{bmatrix} X \\ \cdots \\ x' \end{bmatrix}_{\ell}^{-}$. Observe that $\begin{bmatrix} X \\ \cdots \\ x' \end{bmatrix}_{\ell}^{-} = ((X' \vdots x)_m^{-})'$ and $X_{\ell}^{-} = ((X')_m^{-})'$. Thus the problem is to find $(X' \vdots x)_m^{-}$

given $(X')^-_m$. If $X'(X')^-_m x = x$ then one can use theorem 3(iv) to compute $(X' : x)^-_m$. Otherwise, theorem 1(iii) can be used to compute $(X' : x)^-_m$. Once this is computed the rest of the analysis is easy.

The case where a new parameter is added to the model can be dealt with similarly. This is illustrated by an example in section 2.7.

We now consider

Problem II : Consider the same set up as in (2.6.1). Suppose the model is overspecified and the last component of β is redundant. It is desired to obtain the revised estimates deleting the last component of β and the corresponding (last) column in X.

Let us write $X = (Z : x)$ and $\beta' = (\gamma : \beta_m)$ where β_m is the last component of β and x is the last column of X. The problem is to compute Z^-_ℓ using $(Z : x)^-_\ell$. Let $(Z : x)^-_\ell = (G' : b)'$. Check whether $x'Z = o$. If so using theorem 4(iv), we note that $G = Z^-_\ell$. Otherwise check whether $b'x \neq 1$. If so, use theorem 7(iv) to obtain Z^-_ℓ from X^-_ℓ. If $b'x = 1$, check whether $b'Z = 0'$. If so, use theorem 8(iii) to compute Z^-_ℓ. Otherwise use either theorem 5 or theorem 6(iii). Observe that once Z_ℓ^- is computed the rest of the analysis is easy. The case where an observation is dropped and the revised estimates are to be obtained can be dealt with in a similar way. This again is illustrated by an example in section 2.7.

Now consider the same model as in (2.6.1) but let $D(y) = M\sigma^2$ where M is a known p.d. matrix. Then

$$\hat{\beta} = X^-_{\ell(M^{-1})}\, y \quad \text{and} \quad \hat{\sigma}^2 = \frac{R_o^2}{n - r} \quad \text{where} \quad r = R(X) \quad \text{and}$$

$$R_o^2 = y'M^{-1}y - y'M^{-1}X\hat{\beta} .$$

Let us now consider the problems considered in the earlier case. We shall indicate how one can use the results of section 2.4 to solve these problems. In problem I, we consider only the case where a new observation which is uncorrelated with the previous observations is given. Thus one has to compute

$$\begin{bmatrix} X \\ \cdots \\ x' \end{bmatrix}^-_{\ell(M_1^{-1})} \quad \text{where} \quad M_1 = \begin{bmatrix} M & 0 \\ 0 & \delta \end{bmatrix} .$$

But it is well-known (see Rao and Mitra, 1970) that

$$\begin{bmatrix} X \\ \cdots \\ x' \end{bmatrix}^-_{\ell(M_1^{-1})} = [(X' : x)^-_{m(M_1)}]' \quad \text{and} \quad X^-_{\ell(M^{-1})} = [(X')^-_{m(M)}]' .$$

Now one can use theorem 11 or 12 to obtain $(X' : x)^{-}_{m(M_1)}$ given $(X')^{-}_{m(M_1)}$ according as $x \in \mathcal{M}(X')$ or not. One may also remember that $\mathrm{Var}(z) = \delta\sigma^2 \Rightarrow \mathrm{Var}\left(\frac{z}{\sqrt{\delta}}\right) = \sigma^2$ (where $\delta \neq 0$). Problem II, in this case, can be solved in an exactly similar manner to the previous case and the theorems to be used are given in section 2.4 and do not need any further elaboration.

In either of the two models if one wishes to compute the estimates of the standard errors of estimates of estimable linear parametric functions one should know $(X'M^{-1}X)^{-}$ where $D(y) = M\sigma^2$. This is easily computed from $X^{-}_{\ell(M^{-1})}$ as the following theorem of Rao and Mitra (1970) (See theorem 2(a) of Chapter 1) suggests '' Let $G = X^{-}$. Then $G = X^{-}_{\ell(M^{-1})}$ if and only if $GMG' = (X'M^{-1}X)^{-}$. ''

2.7 Numerical illustration.

Given below are

(i) an observation vector y on a vector Y of random variables obeying a linear model $E(Y) = X\beta$ and $D(Y) = \sigma^2 I$,

(ii) the matrix X of the linear model,

(iii) X_ℓ^- , a least squares g-inverse of X,

(iv) $\hat{\beta} = X_\ell^- y$, a least squares solution for the unknown parametric vector β,

(v) R_o^2, the residual sum of squares

and

(vi) $\hat{\sigma}^2$, an unbiased estimate of σ^2 based on $5-2 = 3$ d.f.

$$X = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} ; \qquad y = \begin{bmatrix} 5.6 \\ 4.2 \\ 9.6 \\ 2.4 \\ 2.4 \end{bmatrix}$$

$$X^- = \begin{bmatrix} \frac{5}{24} & -\frac{1}{12} & \frac{1}{8} & \frac{7}{24} & -\frac{3}{8} \\ -\frac{1}{8} & \frac{1}{4} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} 1.8167 \\ 2.1500 \\ 0.0 \end{bmatrix}$$

$$R_o^2 = y^*y - y^*X\hat{\beta} = 0.51 \quad \text{and} \quad \hat{\sigma}^2 = \frac{R_o^2}{5-2} = 0.17$$

2.7.1 Suppose it is decided to rework the least squares solution excluding the last observation, 2.4.

$$\text{Let } A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad a = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Notice that we are required to work out A_ℓ^-, a least squares g-inverse of A based on X_ℓ^- where $X = \begin{bmatrix} A \\ \cdots \\ a^* \end{bmatrix}$.

$$\text{Let } G^* = \begin{bmatrix} \frac{5}{24} & -\frac{1}{12} & \frac{1}{8} & \frac{7}{24} \\ -\frac{1}{8} & \frac{1}{4} & \frac{1}{8} & -\frac{3}{8} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{and } b^* = \left(-\frac{3}{8} \quad \frac{5}{8} \quad 0 \right)$$

$$\text{Thus } \begin{bmatrix} A \\ \cdots \\ a^* \end{bmatrix}_\ell^- = (G^* : b).$$

Since conjugate transpose of a least squares g-inverse of a matrix is a minimum norm g-inverse of the conjugate transpose, we have

$$(A^* : a)_m^- = \begin{bmatrix} G \\ \cdots \\ b^* \end{bmatrix}$$

Further $b^*a = \frac{5}{8} \neq 1$. Hence applying Theorem 7(iv), we have

$$(A^*)_m^- = G\left(I + \frac{ab^*}{1-b^*a}\right) = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 0 \\ \frac{2}{3} & -1 & 0 \end{bmatrix}$$

Hence $$A_\ell^- = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & 0 & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The revised estimate of β is

$$\hat{\beta} = \begin{bmatrix} 2.0667 \\ 1.7333 \\ 0 \end{bmatrix}$$

$R_0^2 = 0.35$ and $\hat{\sigma}^2 = \frac{R_0^2}{4-2} = 0.175.$

2.7.2 Suppose with the same data as in 2.7.1 we are required to fit a revised model $E(Y) = (X \vdots x) \begin{bmatrix} \beta \\ \cdots \\ \alpha \end{bmatrix}$ where $x^* = (\ 0 \quad 0 \quad 1 \quad 0 \quad 0)$. Then we proceed as follows.

We shall compute $(X \vdots x)_\ell^-$ based on X_ℓ^-. Since $XX^- x \neq x$ it follows that $x \notin \mathcal{M}(X)$. Applying Theorem 1(iv), we have

$$(X \vdots x)_\ell^- = \begin{bmatrix} X_\ell^- - db^* \\ b^* \end{bmatrix}$$

where $d = X_\ell^- x$, $c = (I - XX_\ell^-)x = x - Xd$ and $b = c/c^*x$.

Thus, $d^* = (\frac{1}{8} \ \frac{1}{8} \ 0)$, $c^* = (-\frac{3}{8} \ -\frac{1}{4} \ \frac{3}{8} - \frac{1}{8} . -\frac{1}{8})$ and

$b^* = (-1 \ -\frac{2}{3} \ 1 \ -\frac{1}{3} \ -\frac{1}{3})$.

Hence

$$(X : X)_{\ell}^{-} = \begin{bmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{3} & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -\frac{2}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} .$$

The revised estimate of parametric vector

$$\hat{\begin{bmatrix} \beta \\ \cdots \\ \alpha \end{bmatrix}} = \begin{bmatrix} 1.8667 \\ 2.2000 \\ 0.0 \\ -0.4000 \end{bmatrix}$$

The residual sum of squares, $R_o^2 = 0.45$ and $\hat{\sigma}^2 = \frac{R_o^2}{5-3} = 0.225.$

2.8 <u>An extension of a theorem of Rohde on g-inverses of partitioned matrices.</u>

Rohde (1965) proved the following theorem on g-inverses of partitioned matrices.

<u>Theorem 16 (Rohde)</u> : Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (X_1 : X_2)^*(X_1 : X_2)$

and $G = \begin{bmatrix} A^- + A^-BF^-CA^- & -A^-BF^- \\ -F^-CA^- & F^- \end{bmatrix}$. where

$F = D - CA^-B$. Then the following hold :

(i) $G = M^-$.

(ii) If A^- and F^- in G are replaced by A_r^- and F_r^- then $G = M_r^-$.

(iii) If F is nonsingular and A^- and F^- in G are replaced by A^+ and F^{-1} then G is indeed M^+.

In this section we extend this result. We pose the following question and provide a solution to it.

Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and A^- be a g-inverse of A.
Let $F = D - CA^-B$ and

$$G = \begin{bmatrix} A^- + A^-BF^-CA^- & -A^-BF^- \\ -F^-CA^- & F^- \end{bmatrix} \quad \text{where } F^- \text{ is} \tag{2.8.1}$$

any g-inverse of F.

What are the necessary and sufficient conditions under which G is a g-inverse of M ?

The answer is contained in the following

Theorem 17 : Consider the set up as in(2.8.1). Then

(a) G is a g-inverse of M if and only if

(i) $\mathcal{M}(C(I - A^{-}A)) \subset \mathcal{M}(F)$

(ii) $\mathcal{M}(((I - AA^{-})B)') \subset \mathcal{M}(F')$

and (iii) $(I - AA^{-})BF^{-}C(I - A^{-}A) = 0.$

(b) If A^{-} and F^{-} in the expressions for F and G are replaced by A_r^{-} and F_r^{-}, then M is always a g-inverse of G (no further conditions being required).

Proof :

Proof of (a):

$$MG = \begin{bmatrix} AA^{-} + AA^{-}BF^{-}CA^{-} - BF^{-}CA^{-} & -AA^{-}BF^{-} + BF^{-} \\ CA^{-} - FF^{-}CA^{-} & FF^{-} \end{bmatrix}$$

and
$$MGM = \begin{bmatrix} A + AA^{-}BF^{-}CA^{-}A - BF^{-}CA^{-}A - AA^{-}BF^{-}C + BF^{-}C & AA^{-}B - AA^{-}BF^{-}F + BF^{-}F \\ CA^{-}A - FF^{-}CA^{-}A + FF^{-}C & D \end{bmatrix}$$

Thus $MGM = M$ if and only if

$$\text{(i)}\ AA^{-}B - AA^{-}BF^{-}F + BF^{-}F = B \iff (I - AA^{-})B$$

$$= (I - AA^{-})\, BF^{-}F \iff \mathcal{M}(((I-AA^{-})B)') \subset \mathcal{M}(F')$$

$$\text{(ii)}\ CA^{-}A - FF^{-}CA^{-}A + FF^{-}C = C \iff \mathcal{M}(C(I - A^{-}A)) \subset \mathcal{M}(F).$$

and $\text{(iii)}\ AA^{-}BF^{-}CA^{-}A - BF^{-}CA^{-}A - AA^{-}BF^{-}C + BF^{-}C = 0 \iff$

$$(I - AA^{-})\, BF^{-}C\, (I-A^{-}A) = 0.$$

This completes the proof of (a). Proof of (b) follows by straightforward verification.

<u>Note</u> : In the presence of conditions (i) and (ii), if the condition (iii) holds for some choice of F^{-}, then it holds for every choice of F^{-}.

Now we give a few interesting special cases where the conditions of theorem 17 indeed hold. We prove

<u>Theorem 18</u>:

Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$

where $\mathcal{M}(C') \subset \mathcal{M}(A')$ and $\mathcal{M}(B) \subset \mathcal{M}(A)$ and consider F and G as defined in theorem 17. Then (i) F is invariant under the choices of g-inverse of A and (ii) G is a g-inverse of M.

Proof : (i) follows trivially.

To prove (ii), observe that

$$\mathcal{M}(C') \subset \mathcal{M}(A') \iff C = CA^{-}A, \text{ for any choice of g-inverse } A^{-} \text{ of } A$$

$$\iff (C - CA^{-}A) = 0$$

and $\mathcal{M}(B) \subset \mathcal{M}(A) \iff (I - AA^{-})B = 0$, for any choice of g-inverse A^{-} of A.

Thus if $\mathcal{M}(B) \subset \mathcal{M}(A)$ and $\mathcal{M}(C') \subset \mathcal{M}(A')$ the conditions of theorem 17 are satisfied. Hence G is a g-inverse of M.

The following corollaries are easy to deduce.

Corollary 1 :

$$\text{If } M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (X_1 : X_2)^* (Y_1 : Y_2)$$

where $R(X_1{}^*Y_1) = R(X_1) = R(Y_1)$, then G as defined in theorem 17 is a g-inverse of M.

Corollary 2 : (Rohde) :

If $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (X_1 : X_2)^* (X_1 : X_2)$

then G as defined in theorem 17 is a g-inverse of M.

Theorem 19 : Let M be as defined in theorem 18 and G as in theorem 17. If in G, A^- and F^- are replaced by A_ℓ^- and F_ℓ^-, then G is M_ℓ^- if and only if $\mathcal{M}(CA^-) \subset \mathcal{M}(F)$. If in G, A^- and F^- are replaced by A_m^- and F_m^- then G is M_m^- if and only if $\mathcal{M}((A^-B)') \subset \mathcal{M}(F')$. If in G, A^- and F^- are replaced by A^+ and F^+, then $G = M^+$ if and only if $\mathcal{M}(CA^+) \subset \mathcal{M}(F)$ and $\mathcal{M}((A^+B)') \subset \mathcal{M}(F')$.

Proof of theorem 19 is straightforward and is therefore omitted.

CHAPTER 3

SIMULTANEOUS REDUCTION OF SEVERAL HERMITIAN FORMS

3.1 Introduction and Summary.

Simultaneous reduction of two hermitian forms one of which is positive definite to diagonal forms by means of a nonsingular linear transformation is well-known (See Rao, 1965). In an interesting paper Mitra and Rao (1968) considered the problem of simultaneous redution of a pair of hermitian forms. They obtained in several classified cases neat necessary and sufficient conditions for simultaneous reduction of a pair of hermitian forms to diagonal forms by means of cogredient and contragredient transformations. They also obtained a necessary and sufficient condition for several hermitian forms to be reduced simultaneously to diagonal forms by a single unitary transformation.

In this chapter we obtain in several cases necessary and sufficient conditions for simultaneous reduction of several hermitian forms to diagonal forms by a single nonsingular linear transformation. These are obtained in sections 3.3 - 3.5. In section 3.2, we give a characterisation of semisimple matrices with real eigen values and show that several semisimple matrices commute pairwise if and only if they can be expressed as polynomials in a common semisimple matrix.

It is interesting to find that the conditions for simultaneous reduction expressed in terms of g-inverses (as given here) make the expressions quite neat and intuitive.

In addition to the general notations cited in introduction to the thesis we use the following notation in this chapter. For a $m \times n$ matrix A of rank r, $A^{\perp}$ denotes a matrix of order $m \times m-r$ and of rank $m-r$ such that $A^* A^{\perp} = 0$.

3.2 Some useful theorems.

In this section we prove two useful theorems in linear algebra which are also of independent interest.

Theorem 1 : Let $A_1, A_2, \ldots, A_k$ be semisimple matrices of the same order. Then $A_1, A_2, \ldots, A_k$ commute pairwise if and only if they can be expressed as polynomials in a common semisimple matrix.

<u>Proof</u> : Proof of 'if' part is trivial. To prove the 'only if' part we proceed as follows.

Since A_i, $i = 1, 2, \ldots, k$ are semisimple matrices it follows that (See theorem 9.5 of Perlis, 1956)

$$A_i = a_{i1} E_{i1} + \ldots + a_{i\ell_i} E_{i\ell_i}, \quad i = 1, 2, \ldots, k$$

where

$$\left.\begin{array}{ll} \text{(i)} & E_{ij} \text{ is idempotent } \forall\, i, j \\ \text{(ii)} & E_{ij} E_{ij'} = 0 \text{ if } j \neq j' \ \forall\, i \\ \text{and (iii)} & \sum_{j=1}^{\ell_i} E_{ij} = I \ \forall\, i \end{array}\right\} \qquad (3.2.1)$$

Further, since $A_1, \ldots, A_k$ commute pairwise it follows that (again see theorem 9.5 of Perlis, 1956)

$$E_{ij} E_{\ell m} = E_{\ell m} E_{ij} \quad \forall\, i, j, \ell, m. \qquad (3.2.2)$$

Define $B = \sum_{i_1, \ldots, i_k} c_{i_1, i_2, \ldots, i_k} E_{1i_1} E_{2i_2} \ldots E_{ki_k}$

where the $\ell_1 \ldots \ell_k$ complex numbers $c_{i_1, i_2, \ldots, i_k}$ are all distinct.

First observe that

$$(E_{1i_1} E_{2i_2} \cdots E_{ki_k})(E_{ij_1} \cdots E_{kj_k}) = 0 \text{ if } i_\ell \neq j_\ell \text{ for}$$

some ℓ. This clearly follows from (3.2.1) and (3.2.2). Again from (3.2.1) and (3.2.2) it follows that

$$(E_{1i_1} \cdots E_{ki_k})(E_{1i_1} \cdots E_{ki_k}) = E_{1i_1} \cdots E_{ki_k}$$

and $\sum_{i_1, \ldots, i_k} E_{1i_1} E_{2i_2} \cdots E_{ki_k} = I.$

Now, an appeal to theorem 9.5 of Perlis (1956) yields that B is semisimple.

Further observe that $\sum_{i_2, \ldots, i_k} E_{2i_2} \cdots E_{ki_k} = I$ and also sum of all such similar products by omitting for example the principal idempotents of A_1 is the identity matrix.

Let $p(\,.\,)$ be a polynomial function with complex coefficients. Then observe that from (3.2.1) and (3.2.2) it follows that

$$p(B) = \sum_{i_1,\ldots,i_k} p(c_{i_1,i_2,\ldots,i_k}) E_{1i_1} E_{2i_2} \cdots E_{ki_k}$$

Now there exist polynomials p_m, $m = 1,2,\ldots, k$ such that

$$p_m(c_{i_1,i_2,\ldots,i_k}) = a_{m\,i_m} \quad \text{for all} \quad m, i_1, i_2, \ldots, i_k.$$

[Lagrange's method can be used for constructing such polynomials.]

Hence

$$p_m(B) = \sum_{i_1,\ldots,i_k} p_m(c_{i_1,\ldots,i_k}) E_{1i_1} \cdots E_{ki_k}$$

$$= \Big(\sum_{i_j \neq i_m} E_{1i_1} \cdots E_{ki_k}\Big)\Big(\sum_{i_m=1}^{\ell_m} a_{m\,i_m} E_{m\,i_m}\Big)$$

$$= A_m \quad \text{for} \quad m = 1,2,\ldots,k.$$

This completes the proof of theorem 1.

Corollary : If $A_1, A_2, \ldots, A_k$ are semisimple matrices which commute pairwise then there exists a nonsingular matrix T such that $T A_i T^{-1}$ is diagonal for $i = 1, 2, \ldots, k$.

We now obtain a characterisation of semisimple matrices with real eigen values. In later sections we come across conditions involving semisimple matrices with real eigen values and thus this theorem may not be quite out of place here. We prove

Theorem 2 : A is semisimple with real eigen values if and only if there exists a positive definite matrix M such that $M A M^{-1} = A^*$.

Proof :

Proof of 'if' part.

Let M be a positive definite matrix such that $M A M^{-1} = A^*$. Since M is positive definite $M = BB^*$ where B^* is a nonsingular matrix. Thus

$$BB^* A B^{*-1} B^{-1} = A^* \Rightarrow B^* A B^{*-1} = B^{-1} A^* B$$

$$\Rightarrow B^* A B^{*-1} \text{ is hermitian } \Rightarrow B^* A B^{*-1} \text{ is semisimple}$$

with real eigen values $\Rightarrow$ A is semisimple with real eigen values.

This completes the proof of 'if' part.

Proof of 'only if' part.

If A is semisimple with real eigen values then there exists a nonsingular matrix B such that $B\,A\,B^{-1} = D$ where D is a real diagonal matrix. Hence $B\,A\,B^{-1} = D = D^* = B^{-1*}A^*B^*$. This in turn implies that $B^*B\,A\,B^{-1}B^{*-1} = A^*$. Clearly B^*B is positive definite. This completes the proof of 'only if' part.

The following interesting result of Mitra (1968) follows immediately as an idempotent matrix is indeed a semisimple matrix with real eigen values.

Corollary (Mitra) : If A is an idempotent matrix, then there exists a positive definite matrix M such that $M\,A\,M^{-1} = A^*$.

We state below a theorem given in Mitra and Rao (1968), for completeness.

Theorem 3 : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of the same order. Then there exists a unitary matrix T such that $T A_i T^*$ is diagonal for $i = 1, 2, \ldots, k$ if and only if $A_1, \ldots, A_k$ commute pairwise.

3.3 Simultaneous reduction when one of the matrices is nonsingular.

We prove

Theorem 4 : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of the same order and let A_1 be nonsingular. Then there exists a nonsingular matrix T such that $T A_i T^*$ is diagonal, $i = 1, 2, \ldots, k$ if and only if

(a) $A_i A_1^{-1}$ is semisimple with real eigen values for $i = 1, \ldots, k$

and (b) $A_i A_1^{-1} A_j = A_j A_1^{-1} A_i$ for all i and j.

Proof : Proof of 'only'if' part is trivial. To prove the 'if' part we proceed as follows.

$$\forall\, i, j, \quad A_i A_1^{-1} A_j = A_j A_1^{-1} A_i \Rightarrow A_1 A_1^{-1}, A_2 A_1^{-1}, \ldots, A_k A_1^{-1}$$

commute pairwise.

Hence (a) and (b), in view of theorem 1 and the corollary after theorem 1, imply that there exists a non-singular matrix M such that for each i, $M A_i A_1^{-1} M^{-1} = D_i$ where D_i is a diagonal matrix with real elements. Now,

$\forall\, i, M A_i A_1^{-1} M^{-1} = D_i \Rightarrow \forall\, i, M A_i M^* = D_i M A_1 M^* \Rightarrow$ D_i and $M A_1 M^*$ commute for all i.

Also observe that $D_1, \ldots, D_k$ being diagonal matrices commute pairwise.

Hence by theorem 3, there exists a unitary matrix L such that $L M A_1 M^* L^*$ and $L D_i L^*$ are diagonal for all i.

Thus,

$L M A_i M^* L^* = L D_i L^* L M A_1 M^* L^*$ is diagonal for $i = 1, 2, \ldots, k$.

Put $T = L M$ and observe that T is nonsingular and $T A_i T^*$ is diagonal for all i.

This completes the proof of theorem 4.

<u>Corollary</u> : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of the same order and let A_1 be positive definite. Then there

exists a nonsingular matrix T such that $T A_i T^*$ is diagonal for all i if and only if $A_i A_1^{-1} A_j = A_j A_1^{-1} A_i$ for all i, j.

Proof : Corollary follows trivially from theorem 4 once it is observed that $A_i A_1^{-1}$ is semisimple with real eigen values (being similar to A_i) if A_i is hermitian and A_1 is positive definite.

We now state a theorem analogous to the theorem corresponding to simultaneous reduction of a pair of hermitian matrices by contragredient transformations (See Mitra and Rao, 1968). The proof of this theorem follows on similar lines to that of theorem 4 and is omitted.

Theorem 5 : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of the same order and let A_1 be nonsingular. Then there exists a nonsingular matrix T such that $T A_1 T^*$ and $(T^*)^{-1} A_i T^{-1}$ are diagonal for each i if and only if

(a) $A_i A_1$ is semisimple with real eigen values for $i = 2, \ldots, k$ and (b) $A_i A_1 A_j = A_j A_1 A_i$ for $i, j = 2, \ldots, k$.

3.4 Simultaneous reduction of several arbitrary hermitian forms.

We prove

Theorem 6 : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of order $n \times n$ such that $\mathcal{M}(A_i) \subset \mathcal{M}(A_1)$ for $i = 2,\ldots,k$. Then there exists a nonsingular matrix T such that $T A_i T^*$ is diagonal for all i if and only if

(a) $A_i A_1^-$ is semisimple with real eigen values for all i,

and (b) $A_i A_1^- A_j = A_j A_1^- A_i$ for all i, j.

where A_1^- is some g-inverse of A_1.

Proof :

Proof of 'if' part :

(a) and (b), in view of theorem 1 and the corollary after theorem 1, imply that there exists a nonsingular matrix M such that $M A_i A_1^- M^{-1} = D_i$ where D_i is a real diagonal matrix.

$\mathcal{M}(A_i) \subset \mathcal{M}(A_1)$ and $M A_i A_1^- M^{-1} = D_i \Rightarrow M A_i M^* = D_i M A_1 M^* \Rightarrow D_1,\ldots,D_k$ and $M A_1 M^*$ commute pairwise (since $D_1,\ldots,D_k$ are diagonal they commute pairwise).

The rest of the proof of 'if' part follows on the same lines as that of theorem 4.

Proof of 'only if' part.

Let $R(A_1) = r$ and without loss of generality let

$$T\ A_1\ T^* = \begin{bmatrix} \Delta_1 & 0 \\ 0 & 0 \end{bmatrix}$$ where Δ_1 is a nonsingular diagonal matrix of order $r \times r$.

Partition $T^* = (T_1^* \vdots T_2^*)$ where T_1^* is of order $n \times r$. Observe that $A_1\ T_2^* = 0$ which in turn implies that $A_i\ T_2^* = 0$ for $i = 2,\ldots,k$. (This is so because $\mathcal{M}(A_i) \subset \mathcal{M}(A_1)$.) Hence $T\ A_i\ T^* = \begin{bmatrix} \Delta_i & 0 \\ 0 & 0 \end{bmatrix}$ where Δ_i is a diagonal matrix of order $r \times r$.

Clearly, $A_1^- = T^* \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} T$ is a g-inverse of A_1.

Now,

$$A_i A_1^- = T^{-1} \begin{bmatrix} \Delta_i & 0 \\ 0 & 0 \end{bmatrix} T^{*-1}\ T^* \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} T$$

$$= T^{-1} \begin{bmatrix} \Delta_i\ \Delta_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} T$$

which is clearly semisimple with real eigen values.

It is easy to check that $A_i A_1^- A_j = A_j A_1^- A_i$ for all i, j.

Note : Theorem 4 can be deduced from theorem 6.

We state below a few interesting corollaries which are easy to prove.

Corollary 1 : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices such that $A = \sum_{i=1}^{k} A_i$ is nonnegative definite and $\mathcal{M}(A_i) \subset \mathcal{M}(A)$ for $i = 1, 2, \ldots, k$. Then there exists a nonsingular matrix T such that $T A_i T^*$ is diagonal for $i = 1, 2, \ldots, k$ if and only if $A_i A^- A_j = A_j A^- A_i$ for all i, j where A^- is any g-inverse of A.

Corollary 2 : Let $A_1, A_2, \ldots, A_k$ be n.n.d. (non-negative definite) matrices of the same order. Then there exists a nonsingular matrix T such that $T A_i T^*$ is diagonal for each i, if and only if $A_i A^- A_j = A_j A^- A_i$ for all i, j where $A = \sum_{i=1}^{k} A_i$ and A^- is any g-inverse of A.

We now give a sufficient condition for hermitian matrices $A_1, A_2, \ldots, A_k$ to be reduced simultaneously to diagonal forms by a nonsingular linear transformation. We prove

<u>Theorem 7</u> : Let $A_1, A_2, \ldots, A_k$ be hermitian matrices of the same order. Then there exists a nonsingular matrix T such that $T\ A_i\ T^*$ is diagonal if

(a) $$R(NA) = R(N\ A\ N^*) = R\left(\sum_{i=1}^{k} N\ A_i\right)$$

where $A = \sum_{i=1}^{k} A_i$ and $N^* = A_1^{\perp}$,

(b) $\Gamma_i\ A_1^-$ is semisimple with real eigen values

where $\Gamma_i = A_i - A\ N^*(N\ A\ N^*)^-\ N\ A_i$ and A_1^- is some g-inverse of A_1,

(c) $\Gamma_i\ A_1^-\ \Gamma_j = \Gamma_j\ A_1^-\ \Gamma_i$ for $i, j = 1, 2, \ldots, k$,

(d) $N\ A_i\ N^*\ (N\ A\ N^*)^-$ is semisimple with real eigen values

and (e) $$N\ A_i N^*(NAN^*)^- N\ A_j N^*$$
$$= N\ A_j N^*(NAN^*)^-\ N\ A_i\ N^*$$
for $i, j = 2, \ldots, k$.

<u>Proof</u> : First observe that $\Gamma_i \; N^* = 0$ for $i=1,2,\ldots,k$. Hence $\Gamma_i = J_i \; A_1$ and $\Gamma_i \; A_1^- = \Gamma_i \; A_{1r}^-$ where J_i is some matrix and A_{1r}^- is a reflexive hermitian g-inverse of A_1, for $i = 1,2,\ldots,k$. Let $A_1 = C \; D \; C^*$ where $C^*C = I_r$ and D is a nonsingular diagonal matrix with real diagonal elements. Let $A_{1r}^- = Y \; \Lambda \; Y^*$ where $Y^*Y = I_r$ and Λ is a nonsingular diagonal matrix. Then $A_{1r}^- = Z \; D^{-1} \; Z^*$ where $Z = Y(C^* \; Y)^{-1}$. Clearly $C^*Z = I_r$. Now consider

$$S = \begin{bmatrix} Z^* - GN \\ N \end{bmatrix}$$

where $G = Z^*A \; N^*(N \; A \; N^*)^-$.

Observe that $(Z^* - GN) \; A_i = Z^* \; \Gamma_i$, $i = 1,2,\ldots,k$.

Thus $$S \; A_i \; S^* = \begin{bmatrix} E_i & 0 \\ 0 & N \; A_i \; N^* \end{bmatrix}$$

where $E_i = Z^* \; \Gamma_i \; Z$ for $i = 2,\ldots,k$

and $$S \; A_1 \; S^* = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

Further for each i, $E_i D^{-1}$ is semisimple with real eigen values as $\Gamma_i A_1^-$ is semisimple with real eigen values. Also $\Gamma_i A_1^- \Gamma_j = \Gamma_j A_1^- \Gamma_i$ for all $i, j \Rightarrow E_i D^{-1} E_j = E_j D^{-1} E_i$ for all i, j. Hence by theorem 4 it follows that there exists a nonsingular matrix L such that $L E_i L^*$ is diagonal for $i=2,\ldots,k$ and $L D L^*$ is diagonal. Further (a), (d) and (e) together with theorem 6 imply that there exists a nonsingular matrix M such that $M N A_i N^* M^*$ is diagonal for $i = 2,\ldots,k$.

Let $T = \begin{bmatrix} L & 0 \\ 0 & M \end{bmatrix} S$. Then clearly $T A_i T^*$ is diagonal for $i = 1,2,\ldots,k$. Further T is nonsingular as L, M and S are nonsingular.

This completes the proof of theorem 7.

3.5 Concluding remarks.

The conditions stated in theorem 7 are not necessary. This is exhibited in the following example. Take $k = 3$ and

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ and } A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Observe that condition (a) is violated.

CHAPTER 4

COMPUTATIONAL METHODS OF GENERALIZED INVERSES

4.1 Introduction and Summary.

There are several methods for computing the Moore-Penrose (M-P) inverse (Penrose, 1956 ; Rao, 1965; Ben Israel and Wersan, 1963; Boot, 1963; Pyle, 1964; Golub and Kahan, 1965; Ben Israel, 1965, 1966 to mention some.). To compute A^-, A_ℓ^- or A_m^- one need not compute A^+ as has been pointed out by Rao (1965, 1967). In fact, to compute any of A^-, A_m^-, A_ℓ^-, A^+, it is just sufficient to have an algorithm to compute a generalized inverse of a matrix (See section 4.3). In this chapter we present two computational methods for computing a generalized inverse of a matrix.

The first method, based on reduction of a square matrix to Hermite Canonical Form (HCF) by elementary row operations, is presented in sections 4.2 and 4.3. Here we give a simple and usual pivotal condensation-like algorithm to reduce a square matrix to HCF. As Rao (1965) observed, if $H = BA$ is in HCF and B is nonsingular then $B = A^-$. As is well-known the Gaussian elimination type algorithms are not very stable and this method is no exception. However, this is an elegant and simple method to compute a g-inverse.

The second method is based on reducing an arbitrary matrix to bidiagonal form by orthogonal transformations (Householder type), computing a g-inverse of the bidiagonal matrix and adjusting the g-inverse for the orthogonal transformations made in the beginning. This is described in detail in sections 4.4 and 4.5. This method is the same as that of Golub and Kahan upto the reduction to bidiagonal form. To get M-P inverse they obtain the singular value decomposition of the bidiagonal matrix using a Q-R type algorithm (for QR algorithm, see Francis, 1961, 1962), obtain M-P inverse of the diagonal matrix with singular values in the diagonal (this is very easy as the M-P inverse of a diagonal matrix is very simple to compute) and then adjust for the orthogonal transformations made in the process to get M-P inverse of the given matrix. From the examples that we have worked out on an electronic computer, we find that the method described here is as satisfactory as that of Golub and Kahan. However, only a detailed error analysis (which has not so far been attempted very systematically for g-inverse computations except for the work of Golub and Kahan) can place the several algorithms in the proper order of merit.

4.2 An algorithm to reduce an arbitrary square matrix to Hermite Canonical Form.

Definition 1: A square matrix $A = (a_{ij})$ is said to be in Hermite Canonical Form (HCF) if (i) $a_{ij} = 0$ whenever $i > j$, (ii) a_{ii} is either 0 or 1 for each i, (iii) $a_{ij} = 0$ for all $i \neq j$ if $a_{jj} = 1$ and (iv) $a_{ij} = 0$ for all j if $a_{ii} = 0$.

We give below an algorithm to reduce a given square matrix A to HCF by elementary row operations on A.

Algorithm : Let A be a square matrix of order $n \times n$. The algorithm consists of n sweep-outs. For notational simplicity the (i, j)-th element of the resultant matrix after each sweep-out is denoted by a_{ij}. The detailed steps are given below.

Step 0: Set $i = 1$

Step 1: Check whether $a_{ii} = 0$. If so, go to step 2. Otherwise go to step 5.

Step 2 : Check whether $a_{ji} = 0$ for all j such that $i < j \leq n$. If so, set $m = 1$ and go to step 3. Otherwise let k be the smallest integer $(i < k \leq n)$ such that $a_{ki} \neq 0$. Go to step 4.

Step 3 : Check whether $a_{ji} = 0$ for all j such that $m \leq j < i$. If so go to step 6. If not let k be the smallest integer $(m \leq k < i)$ such that $a_{ki} \neq 0$. Check whether $a_{kk} = 0$. If so go to step 4. Otherwise let $m = k + 1$. Check whether $m = i$. If so, go to step 6. Otherwise start step 3 again.

Step 4 : Interchange i-th and k-th rows. Go to step 5.

Step 5 : Perform sweep-out to make (i,i)-th element unity and all other elements in i-th column zero. Go to step 6.

Step 6 : Check whether $i = n$. If so the process is complete. Otherwise step up i by 1 and go to step 1.

Observe that we have used only elementary row operations all through and that this algorithm indeed reduces a square matrix to HCF. Notice that this algorithm is very similar to the usual Gaussian elimination method for computing the inverse when the matrix under consideration is nonsingular.

4.3 Use of the elimination algorithm in section 4.2 to compute various types of g-inverses.

Let A be a given square matrix. Apply on A the algorithm given in section 4.2. Let B be the matrix which is the product of the matrices corresponding to the elementary row operations made (product taken in the proper order.). Hence $BA = H$ where H is in HCF. Then clearly (See Rao, 1965), B is a g-inverse of maximum rank of A.

If A is a $m \times n$ matrix where $m < n$ then the square matrix $C = \begin{bmatrix} A \\ \cdots \\ 0 \end{bmatrix}$ is constructed. Let D be a g-inverse of C. Delete the last $n - m$ columns in D to get a g-inverse of A. Observe that this is a g-inverse of maximum rank. The case where $m > n$ can be disposed off similarly.

The matrix H itself is also of considerable interest in several applications. So one can apply the algorithm on $A : I$ so as to reduce A to HCF. As a result of the application of the algorithm in place of A and I one gets H and B respectively.

Rao (1967) gave the following expressions for A_ℓ^-, A_m^- and A^+.

$$A_m^- = A'(A\,A')^- \tag{4.3.1}$$

$$A_\ell^- = (A'A)^-\,A' \tag{4.3.2}$$

and $$A^+ = (A'A)_m^-\,A' = A'(A\,A')_\ell^- \tag{4.3.3}$$

for any choices of g-inverses occuring in the expressions on the right hand side in (4.3.1) - (4.3.3).

Hence it follows that

$$A^+ = (A'A)_m^-\,A' = A'A\,(A'A\,A'A)^-\,A' \tag{4.3.4}$$

$$= A'(A\,A')_\ell^- = A'\,(A\,A'A\,A')^-\,A\,A' \tag{4.3.5}$$

To compute A^+ one can use the formula (4.3.4) when $m \geq n$ and the formula (4.3.5) when $m < n$.

Mitra (1968a) gave an equivalent and more compact formula for A^+ as

$$A^+ = A'(A'A\,A')^- A' . \qquad (4.3.6)$$

It may be noted that the formulae (4.3.4) and (4.3.5) can also be derived from (4.3.6).

However for computational purposes (4.3.4) and (4.3.5) are more useful as for example the expressions for which the g-inverses are to be computed are nnd. If C is nnd, in applying the algorithm of the previous section on C, one never needs step 3 as a result of which some checking can be avoided. The previous statement is easy to prove and hence the proof is omitted.

4.4 g-inverse of a bidiagonal matrix.

It is well-known (Golub and Kahan, 1965) that any $m \times n$ matrix can be reduced by Householder's (orthogonal) transformations to bidiagonal form. In this section we exhibit a g-inverse of a block bidiagonal matrix of a particular type. In the next section we shall show that any bidiagonal matrix can be brought to the type, for which a g-inverse is exhibited in this section, by simple elementary operations whence one can compute a g-inverse of an arbitrary matrix using this method. We now define a block bidiagonal matrix.

Definition 2 : A matrix A of order $M \times N$ is said to be a block bidiagonal matrix if it is of the form

$$\begin{bmatrix} A_{11} & A_{12} & 0 & \cdots & 0 \\ 0 & A_{22} & A_{23} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & A_{nn} \end{bmatrix} \qquad (4.4.1)$$

where the matrix in (i, j)-th block is of order $m_i \times n_j$ where $\sum_{i=1}^{n} m_i = M$ and $\sum_{j=1}^{n} n_j = N$. In particular, if each block in the above matrix is an element of the field (i.e. a scalar), the matrix is called a bidiagonal matrix.

Let A be a matrix as defined in definition 2.

Define

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ 0 & B_{22} & \cdots & B_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & B_{nn} \end{bmatrix}$$

where $B_{ii} = A^-_{ii}$ for $i = 1, 2, \ldots, n$

$$B_{ij} = -B_{i,j-1} A_{j-1,j} B_{jj}$$

for each i, $j = i+1, \ldots, n$

and for $i = 1, \ldots, n-1$ (4.4.2)

We prove

<u>Theorem 1</u> : Let A and B be as defined in (4.4.1) and (4.4.2) respectively. If either $\mathcal{M}(A_{i,i+1}) \subset \mathcal{M}(A_{ii})$ for $i = 1, \ldots, n-1$ or $\mathcal{M}(A'_{i,i+1}) \subset \mathcal{M}(A'_{i+1,i+1})$ for $i = 1, \ldots, n-1$ then B is a g-inverse of A.

<u>Proof</u> : Let us denote the (i, j)-th block of AB and ABA by $(AB)_{ij}$ and $(ABA)_{ij}$ respectively. Clearly,

$$(AB)_{ij} = 0 \quad \text{if} \quad i > j$$

$$(AB)_{ii} = A_{ii} B_{ii} \qquad \text{for} \quad i = 1, 2, \ldots, n$$

and for each $i = 1, \ldots, n-1$,

$$(AB)_{ij} = A_{ii}\, B_{ij} + A_{i,i+1}\, B_{i+1,j}$$

for $j = i+1, \ldots, n.$

Hence,

$$(ABA)_{ij} = 0 \quad \text{if} \quad i > j$$

$$(ABA)_{ii} = A_{ii}\, B_{ii}\, A_{ii} = A_{ii} \quad \text{for} \quad i = 1, 2, \ldots, n$$

and for each $i = 1, 2, \ldots, n-1,$

$$(ABA)_{ij} = (A_{ii}\, B_{i,j-1} + A_{i,i+1}\, B_{i+1,j-1})\, A_{j-1,j} + (A_{ii}\, B_{ij} + A_{i,i+1}\, B_{i+1,j})\, A_{jj} \tag{4.4.3}$$

for $j = i+1, \ldots, n$

where B_{ij} is interpretted as 0 if $i > j.$

$$(ABA)_{i,i+1} = A_{ii}\, B_{ii}\, A_{i,i+1} + A_{ii} B_{i,i+1}\, A_{i+1,i+1} + A_{i,i+1}\, B_{i+1,i+1}\, A_{i+1,i+1}$$

$$= A_{ii}\, B_{ii}\, A_{i,i+1} - A_{ii}\, B_{ii}\, A_{i,i+1}\, B_{i+1,i+1}\, A_{i+1,i+1}$$

$$+ A_{i,i+1}\, B_{i+1,i+1}\, A_{i+1,i+1}$$

If either $\mathcal{M}(A_{i,i+1}) \subset \mathcal{M}(A_{ii})$ or $\mathcal{M}(A'_{i,i+1}) \subset \mathcal{M}(A'_{i+1,i+1})$, then clearly $(ABA)_{i,i+1} = A_{i,i+1}$. Choice of i being arbitrary, it follows that $(A\,B\,A)_{i,i+1} = A_{i,i+1}$ if either $\mathcal{M}(A_{i,i+1}) \subset \mathcal{M}(A_{ii})$ or $\mathcal{M}(A'_{i,i+1}) \subset \mathcal{M}(A'_{i+1,i+1})$.

Once again choose and fix i. Let $j \geq i+2$. Now

$$(A\,B\,A)_{ij} = A_{ii}\,B_{i,j-1}\,A_{j-1,j} + A_{i,i+1}\,B_{i+1,j-1}\,A_{j-1,j}$$

$$- (A_{ii}\,B_{i,j-1}\,A_{j-1,j}\,B_{jj}\,A_{jj} +$$

$$A_{i,i+1}\,B_{i+1,j-1}\,A_{j-1,j}\,B_{jj}\,A_{jj})$$

If $\mathcal{M}(A'_{j-1,j}) \subset \mathcal{M}(A'_{jj})$ then clearly $(A\,B\,A)_{ij} = 0$. Also if $\mathcal{M}(A_{i,i+1}) \subset \mathcal{M}(A_{ii})$ then again $(A\,B\,A)_{ij} = 0$ since under this condition, $A_{ii}\,B_{i\ell} = -A_{i,i+1}\,B_{i+1,\ell}$ for all $\ell \geq i+1$.

Hence $A\,B\,A = A$ if either $\mathcal{M}(A_{i,i+1}) \subset \mathcal{M}(A_{ii})$ for $i = 1,\ldots,n-1$ or $\mathcal{M}(A'_{i,i+1}) \subset \mathcal{M}(A'_{i+1,i+1})$ for $i = 1,\ldots,n-1$.

This completes the proof of theorem 1.

The following useful corollary follows as a special case of theorem 1.

Corollary : If $A = (a_{ij})$ is a bidiagonal matrix of order $n \times n$ then B is a g-inverse of A if either

$a_{ii} = o \Rightarrow a_{i,i+1} = o$ for $i = 1,2,\ldots,n-1$ or

$a_{ii} = o \Rightarrow a_{i-1,i} = o$ for $i = 2,\ldots,n$ where $B = (b_{ij})$

is defined as

$b_{ij} = o$ if $i > j$

$b_{ii} = a_{ii}^{-}$ where $a_{ii}^{-} = \frac{1}{a_{ii}}$ if $a_{ii} \neq o$ and

$a_{ii}^{-} = c_i$ where c_i is any constant

if $a_{ii} = o$

for $i = 1,2,\ldots,n$

and $b_{ij} = - b_{i,j-1}\, a_{j-1,j}\, b_{jj}$ for $j = i+1,\ldots,n$

for each of $i = 1,\ldots,n-1$.

Theorem 2 : Let $X = U\,A\,V'$ be a symmetric matrix where U and V are orthogonal matrices and A, a bidiagonal matrix such that $a_{ii} = o \Rightarrow a_{i,i+1} = o$ for $i = 1,2,\ldots,n-1$. Define $b_{ii} = o$ if $a_{ii} = o$. Otherwise define b_{ij} as in the previous corollary. Then $C = V B U'$ is a g-inverse of X such that $\mathcal{M}(C') \subset \mathcal{M}(X)$.

<u>Proof</u> : We need

<u>Lemma 1</u> : If A and B are as defined in the theorem 2 then $\mathcal{M}(B') \subset \mathcal{M}(A)$.

Proof of lemma 1 is computational and we omit the details.

Now, observe that C is a g-inverse of X. Further since X is symmetric $X = VA'U'$. From lemma 1, it follows that $\exists$ D such that $B = D A'$. Hence $X = V D A'U' = V D V'V A'U' = V D V'X$.

This completes the proof of theorem 2.

The above theorem can be useful in the following situation. Let $Z Z' = X$. Suppose $X = U A V'$ where U, V and A are as defined in theorem 2 and let B be computed as in theorem 2. Then $Z'C = Z^+$. Thus, in such a case, the minimum norm g-inverse of Z that is computed by the above process is indeed Z^+.

4.5 <u>Another algorithm for computing g-inverses.</u>

In this section we present an algorithm to compute a g-inverse of a matrix using the corollary after theorem 1 and the fact that any arbitrary matrix can be reduced to bidiagonal form by Householder's orthogonal transformations. Once we have an algorithm to compute a g-inverse, it can be made use of to compute various types of g-inverses as has already been described in section 4.3.

Algorithm : Let X be the given $m \times n$ matrix $(m \geq n)$.

Step 1 : Reduce X by means of Householder's transformations to bidiagonal form (See Golub and Kahan, 1965). Let $U'XV = A$ where U and V are orthogonal and A is bidiagonal.

Step 2 :

Substep 2.1 : Set $i = 1$. Go to substep 2.2.

Substep 2.2 : Is $a_{ii} = 0$? If yes, go to substep 2.3. Otherwise go to substep 2.6.

Substep 2.3 : Is $a_{i,i+1} = 0$? If yes, go to substep 2.6. Otherwise go to substep 2.4.

Substep 2.4 : Set $\ell_i = 1$. Set $a_{ii} = a_{i,i+1}$. (Replace the value of a_{ii} by that of $a_{i,i+1}$). Go to substep 2.5.

Substep 2.5 : Is $a_{i+1,i+1} = o$? If yes, set $m_i = o$ and go to substep 2.7.
If no, set $m_i = 1$ and

$$p_i = -\frac{a_{i+1,i+1}}{a_{ii}} \quad \text{and} \quad a_{i+1,i+1} = o.$$

Go to substep 2.7.

Substep 2.6 : Set $\ell_i = o$. Go to substep 2.7.

Substep 2.7 : Is $i = n-1$? If yes, go to step 3. Otherwise step up i by 1 and go to substep 2.2.

Step 3 : Compute the $n \times m$ matrix B as follows :

$$b_{ii} = \begin{cases} a_{ii}^{-1} & \text{if } a_{ii} \neq o \\ o & \text{if } a_{ii} = o \end{cases} \qquad \text{for } i = 1,2,\ldots,n$$

$$b_{ij} = -\, b_{i,j-1}\, a_{j-1,j}\, b_{jj} \qquad \text{for } i = 1,2,\ldots,n-1 \text{ and for each } i,\; j = i+1,\ldots,n.$$

and $b_{ij} = o$ whenever $i > j$ or $j > n$.

Go to step 4.

Step 4 :

Substep 4.1 : Set $i=n-1$. Go to substep 4.2.

Substep 4.2 : Is $\ell_i = 1$? If yes, replace ''(i+1)-th row of b'' by ''(i-th row of B) +((i+1)-th row of b)'' go to substep 4.3. Otherwise go to substep 4.4.

Substep 4.3 : Is $m_i = 1$? If yes, replace ''i-th column of B'' by ''(i-th column of B) + p_i((i+1)-th column of B)''. Go to substep 4.4.

Substep 4.4 : Is $i = 1$? If yes, go to step 5. Otherwise step down i by 1 and go to substep 4.2.

Step 5 : Compute $X^- = V\,B\,U'$.

A few interesting remarks are in order.

Remark 1 : In step 2 we make some simple elementary row and column operations on the bidiagonal matrix obtained through step 1 (if necessary) so that we can apply corollary following theorem 1 to get a g-inverse of this matrix. In step 3 we actually compute the g-inverse of the bidiagonal matrix as described in the corollary. In step 4 we adjust the g-inverse for the elementary operations made in step 2. In step 5, we adjust the g-inverse for the orthogonal transformations made in step 1.

Remark 2 : At the end of step 1, $R(X) = R(A)$. Observe that step 2 does not alter the rank of A. In step 3 we compute a reflexive g-inverse of A. In fact, we can compute a g-inverse of specified rank. For example if $R(A) = r < n$, to obtain a g-inverse of rank $r + p$ (where $0 \leq p \leq n-r$) make $b_{ii} = c$ where $c \neq 0$ when $a_{ii} = 0$ for p of the zero diagonal elements in A and for the rest $n-r-p$ zeros in diagonal make the corresponding diagonal elements in the g-inverse (of bidiagonal matrix) zero. It is easy to check that in this case we get a g-inverse of rank $r + p$. (Observe that $R(B) = R(C)$.).

4.6 Numerical illustration.

In this section, we give two examples, the first using the elimination algorithm given in section 4.3 and the other using the second algorithm in the previous section.

$$A = \begin{bmatrix} -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 1 & -1 & -3 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & -2 \end{bmatrix}$$

$$A^{-} = \begin{bmatrix} -1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ -1.5000 & 1.5000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 \\ -0.5000 & 0.5000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

$$A_m^{-} = \begin{bmatrix} -0.8333 & 0.1667 & 1.5000 & 1.0000 & 0.0000 & 1.0000 \\ -0.3333 & 0.6667 & -1.0000 & -1.0000 & 0.0000 & 0.0000 \\ 1.1667 & -0.8333 & -0.5000 & 0.0000 & 0.0000 & -1.0000 \\ -0.5000 & 0.5000 & 1.5000 & 1.0000 & 0.0000 & -2.0000 \end{bmatrix}$$

$$A_{\ell}^{-} = \begin{bmatrix} -0.6250 & 1.0000 & 0.1250 & -0.1250 & 0.7500 & 0.6250 \\ -0.5000 & 1.5000 & 0.0000 & 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.1250 & 0.5000 & 0.1250 & -0.1250 & 0.2500 & 0.1250 \end{bmatrix}$$

$$A^{+} = \begin{bmatrix} -0.2500 & 0.1667 & 0.0833 & -0.0833 & 0.3333 & 0.2500 \\ -0.1250 & 0.6667 & -0.0417 & 0.0417 & 0.0833 & 0.1250 \\ 0.3750 & -0.8333 & -0.0417 & 0.0417 & -0.4167 & -0.3750 \\ -0.1250 & 0.5000 & 0.1250 & -0.1250 & 0.2500 & 0.1250 \end{bmatrix}$$

<u>Example 2</u> : The following example is worked out on IBM 1401 EDPM using a program in FORTRAN II with mantissa 15 and modulus 5. The results are represented in E mode.

$$A = \begin{bmatrix} 22 & 10 & 2 & 3 & 7 \\ 14 & 7 & 10 & 0 & 8 \\ -1 & 13 & -1 & -11 & 3 \\ -3 & -2 & 13 & -2 & 4 \\ 9 & 8 & 1 & -2 & 4 \\ 9 & 1 & -7 & 5 & -1 \\ 2 & -6 & 6 & 5 & 1 \\ 4 & 5 & 0 & -2 & 2 \end{bmatrix}$$

$$(A^{-})' = \begin{bmatrix}
0.65162037E-01 & -0.41666667E-01 & 0.33564815E-02 & -0.63425926E-01 & -0.39583333E-01 \\
-0.54150507E-00 & 0.58117138E+00 & -0.47820580E-01 & 0.88833421E-00 & 0.55279088E-00 \\
0.56639806E+00 & -0.57959906E+00 & 0.67773847E-01 & -0.95918501E+00 & -0.56485849E-00 \\
-0.95583508E-00 & 0.99135220E+00 & -0.83874913E-01 & 0.15237072E+01 & 0.94449686E-00 \\
0.16106744E-00 & -0.15369497E-00 & 0.18404962E-01 & -0.25593117E-00 & -0.15007862E-00 \\
0.45407276E-00 & -0.46481918E+00 & 0.34431778E-01 & -0.69657145E-00 & -0.43954403E-00 \\
-0.65859320E-00 & 0.68238994E+00 & -0.67018256E-01 & 0.10826957E+01 & 0.65640723E-00 \\
0.15683089E-00 & -0.15585692E-00 & 0.17575122E-01 & -0.25555992E-00 & -0.15145440E-00
\end{bmatrix}$$

$$
(A_{R}^{-})' = \begin{bmatrix}
0.40987089E-01 & -0.16256503E-01 & 0.35667045E-03 & -0.24428382E-01 & -0.15761305E-01 \\
-0.20283880E-00 & 0.22520101E-00 & -0.57962992E-02 & 0.34201856E-00 & 0.21906865E-00 \\
0.19921696E-00 & -0.19365688E-00 & 0.22211229E-01 & -0.36687098E-00 & -0.20303770E-00 \\
-0.37476836E-00 & 0.38059594E-00 & -0.11771743E-01 & 0.58636599E-00 & 0.37191286E-00 \\
0.63065678E-01 & -0.50685822E-01 & 0.62441595E-02 & -0.97840741E-01 & -0.53507542E-01 \\
0.18879932E-00 & -0.18599162E-00 & 0.15146360E-02 & -0.26864860E-00 & -0.17814319E-00 \\
-0.24501516E-00 & 0.24768017E-00 & -0.15698353E-01 & 0.41553692E-00 & 0.24886682E-00 \\
0.59046124E-01 & -0.53075853E-01 & 0.54412466E-02 & -0.97819539E-01 & -0.55097154E-01
\end{bmatrix}
$$

$$
(A_m^-)' = \begin{bmatrix}
0.21129808E-01 & 0.46153846E-02 & -0.21073718E-02 & 0.76041667E-02 & 0.38060897E-02 \\
0.93108974E-02 & 0.22115385E-02 & 0.20528846E-01 & -0.20833333E-03 & 0.10016026E-01 \\
-0.11097756E-01 & 0.27403846E-01 & -0.38862179E-02 & -0.27604167E-01 & 0.42067308E-02 \\
-0.79166667E-02 & -0.50000000E-02 & 0.33750000E-01 & -0.54166667E-02 & 0.10416667E-01 \\
0.55128205E-02 & 0.98076923E-02 & -0.89743590E-03 & -0.50000000E-02 & 0.32051282E-02 \\
0.14318910E-01 & -0.25961538E-02 & -0.20136218E-01 & 0.12812500E-01 & -0.62099359E-02 \\
0.48958333E-02 & -0.15000000E-01 & 0.15312500E-01 & 0.12395833E-01 & 0.26041667E-02 \\
0.15064103E-02 & 0.74038462E-02 & -0.16987179E-02 & -0.50000000E-02 & 0.16025641E-02
\end{bmatrix}
$$

$$(A^{+})' = \begin{bmatrix} 0.21129808E-01 & 0.46153846E-02 & -0.21073718E-02 & 0.76041667E-02 & 0.38060897E-02 \\ 0.93108974E-02 & 0.22115385E-02 & 0.20528846E-01 & -0.20833333E-03 & 0.10016026E-01 \\ -0.11097756E-01 & 0.27403846E-01 & -0.38862179E-02 & -0.27604167E-01 & 0.42067308E-02 \\ -0.79166667E-02 & -0.50000000E-02 & 0.33750000E-01 & -0.54166667E-02 & 0.10416667E-01 \\ 0.55128205E-02 & 0.98076923E-02 & -0.89743590E-03 & -0.50000000E-02 & 0.32051282E-02 \\ 0.14318910E-01 & -0.25961538E-02 & -0.20136218E-01 & 0.12812500E-01 & -0.62099359E-02 \\ 0.48958333E-02 & -0.15000000E-01 & 0.15312500E-01 & 0.12395833E-01 & 0.26041667E-02 \\ 0.15064103E-02 & 0.74038462E-02 & -0.16987179E-02 & -0.50000000E-02 & 0.16025641E-02 \end{bmatrix}$$

Remark : This is an example for which AA' satisfies the conditions of theorem 2 and thus as pointed out in the comment after theorem 2, A_m^- computed by this algorithm is indeed A^+ as is observed above.

R E F E R E N C E S

Ben-Israel, A. (1965) : An iterative method for computing the generalized inverse of an arbitrary matrix. Math. Comp., 19, 452-455.

______ (1966) : A note on an iterative method for generalized inversion of matrices. Math. Comp., 20, 439-440.

Ben-Israel, A. and Charnes, A. (1963) : Contributions to the theory of generalized inverses. SIAM, J. Appl. Math., 11, 667-699.

Ben-Israel, A. and Wersan, S.J. (1963) : An elimination method for computing the generalized inverse of an arbitrary complex matrix. J. Assoc. Comput. Mach., 10, 532-537.

Bhimasankaram, P. (1971) : On generalized inverses of partitioned matrices. (To appear in Sankhyā, Ser. A., 33.)

Bhimasankaram, P. and Mitra, S.K. (1969) : On a theorem of Rao on g-inverses of matrices. Sankhyā, Ser. A, 31, 365-368.

Bjerhammer, A. (1951) : Rectangular reciprocal matrices, with special reference to geodetic calculations. Bull. Geodesique., 52, 188-220.

Boot, J.C.G. (1963) : The computation of the generalized inverse of singular or rectangular matrices. Amer. Math. Monthly, 70, 302-303.

Cline, R. C. (1964) : Representations for the generalized inverse of a partitioned matrix. SIAM, J. Appl. Math., 12, 588-600.

Francis, J.G.F. (1961,'62) : The QR transformation, Parts I and II. Computer, J., 4, 265-271, 332-345.

Golub, G. and Kahan, W. (1965) : Calculating the singular values and the pseudoinverse of a matrix. SIAM, J. Numer. Anal., Ser. B., 2, 205-224.

Greville, T.N.E. (1960) : Some applications of the pseudo-inverse of a matrix. SIAM Rev., 2, 15-22.

________ (1961) : Note on fitting of functions of several independent variables. SIAM, J. Appl. Math., 9, 109-115, 317.

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

Mitra, S.K. (1968a) : On a generalized inverse of a matrix and applications. Sankhyā, Ser. A., 30, 107-114.

Mitra, S.K. (1968b) : A new class of g-inverse of square matrices. Sankhyā, Ser. A., 30, 323-330.

Mitra, S.K. and Bhimasankaram, P. (1970a) : Some results on idempotent matrices and a matrix equation connected with the distribution of quadratic forms. Sankhyā, Ser. A., 32, 353-356.

__________ (1970b) : Generalized inverse of partitioned matrices and recalculation of least squares estimates for data or model changes. R.T.S. Tech. Rep. No. Math-Stat./11/70, Indian Statistical Institute.

Mitra, S.K. and Rao, C.R. (1968) : Simultaneous reduction of a pair of quadratic forms. Sankhyā, Ser. A., 30, 313-322.

Moore, E. H. (1920) : On the reciprocal of the general algebraic matrix. (abstract). Bull. Amer. Math. Soc., 26, 394-395.

__________ (1935) : General Analysis, Part 1. Mem. American Phil. Soc., Philadelphia.

Morris, G.L. and Odell, P.L. (1968) : Common solutions for n matrix equations with applications. J. Assoc. Comput. Mach., 15, 272-274.

Ogasawara, T. and Takahashi, M. (1951) : Independence of quadratic forms in normal system. Journal of Science of Hiroshima University, 15, 1-9.

Penrose, R. (1955) : A generalized inverse for matrices. Proc. Cambridge Philos. Soc., 51, 406-413.

____________(1956) : On best approximate solutions of linear matrix equations. Proc. Cambridge Philos. Soc., 52, 17-19.

Perlis, S. (1952) : Theory of matrices. Addison-Wesley Pub. Co., Cambridge, Mass.

Pringle, R. M. and Rayner, A. A. (1970) : Expressions for generalized inverses of a bordered matrix with application to the theory of constrained linear models. SIAM Rev., 12.

Pyle, L.D. (1964) : Generalized inverse computations using the gradient projection method. J. Assoc. Comput. Mach., 11, 422-428.

Rado, R. (1956) : Note on generalized inverses of matrices. Proc. Cambridge Philos. Soc., 52, 600-601.

Rao, C. R. (1955) : Analysis of dispersion for multiply classified data with unequal numbers in cells. Sankhyā, **15**, 253-280.

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

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

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

Rao, C.R. and Mitra, S.K. (1970) : Generalized inverse of matrices and its applications. John Wiley and Sons, New York. (In press).

Rohde, C.A. (1965) : Generalized inverses of partitioned matrices. SIAM, J. Appl. Math., 13, 1033-1035.

Scroggs, J.E. and Odell, P.L. (1966) : An alternative definition of a pseudoinverse of a matrix. SIAM, J. Appl. Math., 14, 796-810.

Tseng, Y.Y. (1949a) : Generalized inverses of unbounded operators between two unitary spaces. Dokl. Akad. Nauk. SSSR (N.S.), 67, 431-434.

Tseng, Y.Y. (1949b) : Properties and classifications of generalized inverses of closed operators. Dokl. Akad. Nauk. SSSR. (N.S.), 67, 607-610.

____________(1956) : Virtual solutions and general inversions. Uspehi. Mat. Nauk. (N.S.), 11, 213-215.