CALCULUS OF GENERALIZED INVERSES OF MATRICES PART I—GENERAL THEORY

By C. RADHAKRISHNA RAO
Indian Statistical Institute

I. GENERAL THEORY

SUMMARY. Singular square matrices and rectangular matrices do not possess inverses in the regular sense of the term. None-the-less, for certain purposes such as solving consistent linear equations or obtaining least square solutions of inconsistent linear equations, inverses of such matrices can be defined and used in the same way as a regular inverse. The name of generalized inverse (g-inverse) is used in such cases to distinguish it from a regular inverse. The paper shows how a g-inverse can be defined depending on the purpose for which it is used. It also attempts at a classification of g-inverses based on their uses and discusses their interrelationships.

1. Introduction

The principal contributors to the subject of generalized inverses are Moore (1935), Murray and Von Neumann (1936) and Bjerhammer (1961). A systematic account is contained in two papers by Penrose (1955, 1956). In these papers the authors define a g-inverse by a set of conditions and show that there exists a unique matrix satisfying the conditions. A similar notion was also used by Bott and Duffin (1953) under the name of constrained inverse and by Aitken (1934) with a different symbolism.

In 1955, the author (Rao, 1955) constructed what is termed as a pseudoinverse of a singular matrix, which does not satisfy all the conditions of Moore-Penrose g-inverse, but is found useful in the discussion of problems involving solutions of equations with singular square matrices and rectangular matrices. It appears that for many applications it is sufficient to work with a g-inverse satisfying a more general (weaker) definition, although such an inverse may not be unique. Therefore, a general definition of a g-inverse was proposed in another paper (Rao, 1962) and its properties were examined. In the 1962 paper some applications of g-inverse to problems in mathematical statistics are also considered. Since then, a number of papers have appeared exploring further uses of generalized inverses, specially in the discussion of the theory of least squares, extrema of quadratic forms, distribution of quadratic forms etc. Some of these applications have been discussed in some detail in two later publications (Rao, 1965 ; 1966).

The term pseudoinverse is also used in a paper by Greville (1957). Wilkinson (1958) uses the term, 'effective inverse' for inverting normal equations in the least square problem and mentions that the application was suggested to him by A. T. James in a personal communication in 1956. Bose (1959) mentions the use of g-inverses in his lecture notes on Analysis of Variance.

The object of the paper is to provide a calculus of the generalized inverses extending the previous work. Applications to problems in mathematical statistics will be considered in a forthcoming paper (Part II).

We consider different definitions of a g-inverse of a matrix depending on the purpose for which such an inverse is needed. Some definitions, such as the general definition of a g-inverse given in Section 2, are valid for matrices with elements belonging to any field. Others involving concepts such as norm and orthogonal projections are valid only for matrices with real or complex elements or for linear operators in finite dimensional vector spaces furnished with an inner product. In the present communication we consider only matrices with real or complex elements unless otherwise stated.

Notations and some basic results in matrix theory. We need some basic results in matrix theory in the discussion of generalized inverses, which are briefly reviewed in this section.

A matrix is generally denoted by a capital letter such as $A, B, C, \dots$. The conjugate transpose of a matrix A, denoted by A^*, is a matrix obtained from A by interchanging rows and columns and replacing the elements, if they are complex, by their conjugates. Two column vectors X and Y are said to be orthogonal if $X^*Y = 0 = Y^*X$. The following results may be easily verified.

(a) $A^{**} = A$

(b) rank $A =$ rank $A^*A =$ rank $AA^* =$ rank A^*

(c) $A^*A = 0 \Longleftrightarrow A = 0$

(d) Let B be a $p \times m$ matrix of rank m, C be a $n \times q$ matrix of rank n and A be a $m \times n$ matrix. Then $BAC = 0 \Longleftrightarrow A = 0$.

The result (d) follows by multiplying $BAC = 0$ from the left by B^* and from the right C^* and observing that B^*B and CC^* are square matrices with full rank.

The linear manifold generated by the columns of a matrix is represented by $M(A)$. A set containing maximal number of independent vectors orthogonal to the columns of A is represenred by (the columns of) matrix $A^{\perp}$. A matrix P such that, for any given vector X, PX is the orthogonal projection onto a given linear manifold is called projection operator. We have the following results.

(a) Let A be a $m \times n$ matrix.

Then $\quad$ rank $A+$rank $A^{\perp} = m$

$\quad\quad\quad$ rank A^*+rank $A^{*\perp} = n$

(b) P is symmetric (i.e., $P = P^*$) and idempotent, (i.e. $P^2 = P$) $\Longleftrightarrow P$ projects vectors onto $M(P)$.

Let P project vectors on $M(A)$ and choose an arbitrary vector X. Then $PX\epsilon M(A)$ and $(X-PX)\epsilon M(A^{\perp})$. Hence $P^{*}A^{\perp}=0$ and $(I-P)^{*}A=0$. Then

$$P^{*}(I-P)^{*}(A\,|\,A^{\perp})=0\Longrightarrow P^{*}(I-P)^{*}=0 \text{ or } P^{2}=P.$$

i.e., P is idempotent. Further

$$A^{*}(I-P)^{*}(A\,|\,A^{\perp})=0\Longrightarrow A^{*}(I-P)^{*}=0 \text{ or } PA=A.$$

Also,

$$(I-P)\,P^{*}(A\,|\,A^{\perp})=((I-P)A\,|\,0)=0\Longrightarrow P^{*}=PP^{*}.$$

Hence $P=PP^{*}$, i.e., P is symmetric. The relations $P^{*}A^{\perp}=0$ and $PA=A$ show that $M(A)=M(P)$.

2. A general definition of a g-inverse

If A is a non-singular square matrix, then the solution of the linear equation $AX=Y$, where Y is $n\times 1$ column vector, is given by $X=A^{-1}Y$ where A^{-1} is the inverse of A (i.e., $AA^{-1}=A^{-1}A=I$, the identity matrix). We ask the question whether a similar representation of the solution, i.e., in the form $X=BY$, is possible when A is a singular square matrix or a rectangular matrix. If there exists a matrix B such that $X=BY$ is a solution of $AX=Y$ for any Y such that $AX=Y$ is a consistent equation, then B does the same job (or behaves) as the inverse of A and hence can be called a generalized inverse of A. So we give a formal definition of a generalized inverse.

Definition 1 : Consider an $m\times n$ matrix A of any rank. A generalized inverse (or g-inverse) of A is an $n\times m$ matrix, denoted by A^{-}, such that for any vector Y for which $AX=Y$ is a consistent equation, $X=A^{-}Y$ is a solution.

We establish the existence of A^{-} and investigate its properties.

Lemma 2a : *A^{-} exists* $\Longleftrightarrow AA^{-}A=A$.

Suppose A^{-} exists. Choose Y as the i-th column a_i of A. The equation $AX=a_i$ is clearly consistent and hence $X=A^{-}a_i$ is a solution, that is, $AA^{-}a_i=a_i$ for all i which is the same as $AA^{-}A=A$.

Conversely, if A^{-} is such that $AA^{-}A=A$ and $AX=Y$ is consistent, then $AA^{-1}AX=AX$ or $A(A^{-}Y)=Y$. Hence $X=A^{-}Y$ is a solution, which proves Lemma 2a. The result of Lemma 2a shows that we have the following equivalent definition of a g-inverse.

Definition 2 : A g-inverse of A is a matrix A^{-} such that

$$AA^{-}A=A. \qquad \dots \quad (2.1)$$

We show that A^{-} exists although it may not be unique.

Lemma 2b : *A^- exists and rank $A^- \geqslant$ rank A.*

Given a $m \times n$ matrix A, there exist non-singular matrices B, $(m \times m)$, and C, $(n \times n)$, such that $BAC = \Delta$ or $A = B^{-1}\Delta C^{-1}$ where Δ is a diagonal matrix (not necessarily square), i.e., with non-zero elements possible only in the main diagonal and zero elements elsewhere. Define by Δ^- the matrix obtained by replacing the non-zero elements of Δ by their reciprocals and taking the transpose. Then it may be easily seen that $\Delta\Delta^-\Delta = \Delta$. Consider $A^- = C\Delta^-B$ and verify that $AA^-A = B^{-1}\Delta C^{-1} = A$, so that A^- is a generalized inverse. Obviously rank $A^- \geqslant$ rank A.

We shall now consider some results in connection with the solution of consistent linear equations $AX = Y$. We have already seen that $X = A^-Y$ is a solution. But there may be more than one solution in which case we may like to have an algebraic expression for a general solution or for generating all possible solutions. Fortunately, this is possible with the help of a g-inverse alone, as defined in (2.1).

Theorem 2a : *Let A^- be any g-inverse of A and $A^-A = H$. Then*

(a) *$H^2 = H$, that is H is idempotent.*

(b) *$AH = A$ and rank A = rank H = trace H.*

(c) *A general solution of the homogeneous equation $AX = 0$ is $(I-H)Z$ where Z is an arbitrary vector.*

(d) *A general solution of a consistent non-homogeneous equation, $AX = Y$, is*

$$A^-Y+(H-I)Z \qquad \dots \quad (2.2)$$

where Z is an arbitrary vector.

(e) *$Q'X$, where Q is a given vector, has a unique value for all solutions of $AX = Y$, iff*

$$H'Q = Q \qquad \dots \quad (2.3)$$

i.e., *Q is an eigen vector of H.*

(f) *The necessary and sufficient condition that $AX = Y$ is consistent is that $AA^-Y = Y$.*

Proof of (a) : $H^2 = A^-A\,A^-A = A^-(A\,A^-A) = A^-A = H$.

Proof of (b) : $AH = AA^-A = A$ so that rank $A \leqslant$ rank H. Since $A^-A = H$, rank $H \leqslant$ rank A. Hence rank A = rank H = trace H, since by (a) H is an idempotent matrix.

Proof of (c) : $A(I-H)Z = 0$ for any Z and hence $(H-I)Z$ is a solution of $AX = 0$. All solutions can be obtained by varying Z, since

$$\text{rank}\,(I-H) = \text{trace}\,(I-H) = \text{trace}\,I - \text{trace}\,H$$
$$= n - \text{rank}\,A$$

which is the rank of the solution space of $AX = 0$.

Proof of (d) : The result (2.2) follows since a general solution of $AX = Y$ is the sum of a particular solution of $AX = Y$ and a general solution of $AX = 0$.

Proof of (e) :

$$\begin{aligned} Q'X &= Q'[A^-Y+(I-H)Z] \\ &= Q'A^-Y+(Q'-Q'H)Z \\ &= Q'A^-Y, \text{ since } Q' = Q'H \end{aligned}$$

so that the result is the same whatever Z may be.

Proof of (f) : The result is obvious.

Lemma 2c : *The necessary and sufficient condition that*

$$BA^-A = B \qquad \dots \quad (2.4)$$

is that $B = DA$, *where* D *is arbitrary.*

Sufficiency is obvious. To prove the necessity let A be a $m \times n$ matrix of rank r and C be $n-r \times n$ matrix of rank $n-r$ such that $AC^* = 0$. By choice CC^* is a non-singular square matrix. Any given matrix B can be written

$$B = \begin{pmatrix} D & \vdots & E \end{pmatrix} \begin{pmatrix} A \\ \cdots \\ C \end{pmatrix} = DA+EC. \qquad \dots \quad (2.5)$$

From (2.4), we find that

$$BC^* = BA^-AC^* = 0, \text{ since } AC^* = 0. \qquad \dots \quad (2.6)$$

Now multiplying both sides of (2.5) by C^*,

$$BC^* = 0 = DAC^*+ECC^* = ECC^* = 0.$$

Since CC^* is non-singular, $E = 0$ giving the required result.

Theorem 2b : *The following results hold for any g-inverse.*

(a) *One choice of* $(A^*)^-$ *is* $(A^-)^*$.

(b) $A(A^*A)^-A^*A = A$.

(c) $(A^*A)(A^*A)^-A^* = A^*$.

(d) $A(A^*A)^-A^*$ *is symmetrical, idempotent and of rank equal to that of* A.

(e) $A(A^*A)^-A^*$ *is a projection operator which projects vectors onto the subspace spanned by the columns of* A.

(f) *If* A *is idempotent,* $A^- = A$ *is one choice of a g-inverse of* A.

(g) $A^*AA^-A = A^*A \Longleftrightarrow AA^-A = A$.

The result (a) follows from definition. To prove (b) consider

$$\begin{aligned} &[A(A^*A)^-A^*A-A]^*[A(A^*A)^-A^*A-A] \\ &= [(A^*A)^-A^*A-I]^*A^*[A(A^*A)^-A^*A-A] \\ &= [(A^*A)^-A^*A-I][A^*A(A^*A)^-A^*A-A^*A] \end{aligned}$$

which vanishes because the second factor vanishes by definition. Hence $A(A^*A)^-(A^*A) = A$. (c) is proved in a similar way.

To prove (d), first consider $A(A^*A)^-A^*-A[(A^*A)^-]^*A^*$ and multiply by its conjugate transpose. Using the result (c), the product is found to be zero showing that

$$A(A^*A)^-A^* = [A(A^*A)^-A^*]^*.$$

The idempotency and the result on rank are easily established.

By (d), $A(A^*A)^-A^*$ is a projection operator. If X is an arbitrary vector, then obviously $A(A^*A)^-A^*X \epsilon M(A)$. The projection is onto, since rank $(A(A^*A)^-A^*)$ = rank A. Thus (e) is proved. It is easy to establish (f).

(g) is proved by showing that the product of (AA^-A-A) by its transpose is a zero matrix.

Theorem 2c. *Let A_i be $m \times n_i$ matrix of rank r_i, $i = 1, ..., k$. If $\Sigma r_i = m$, then the statements*

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

(b) $I = A_1(A_1^*A_1)^-A_1^*+...+A_k(A_k^*A_k)^-A_k$,

are equivalent where in (b), $(A_i^*A_i)^-$ *is any g-inverse of* $(A_i^*A_i)$.

To prove (a) $\Longrightarrow$ (b), let the right side of (b) be B. Then multiplying both sides of (b) by A_i^* and using (a) we have

$$A_i^*(I-B) = A_i^*-A_i^*A_i(A_i^*A_i)^-A_i^* = 0 \quad \text{... (2.7)}$$

using (c) of Theorem 2b. Let $C = (A_1 \,\vdots\, A_2 \,\vdots\, ...\, A_k)$. Then $C^*(I-B) = 0$ by (2.7). But rank $C = m$. Hence $I-B = 0$.

To prove (b) $\Longrightarrow$ (a) we observe that

$$A_i(A_i^*A_i)^-A_i^* = E_i(E_i^*E_i)^{-1}E_i^*$$

where E_i is the matrix obtained from A_i by retaining only the independent columns and omitting the rest. Thus the order of E_i is $m \times r_i$ with rank equal to r_i and $E_i^*E_i$ admits a true inverse since its rank is full. Then (b) can be written

$$I = E_1(E_1^*E_1)^{-1}E_1^*+ \;...\; +E_k(E_k^*E_k)^{-1}E_k \quad \text{... (2.8)}$$

where $E = (E_1 \,\vdots\, E_2 \,\vdots\, ... \,\vdots\, E_k)$ is of rank m. The equation (2.8) can be written as

$$I = E \begin{pmatrix} (E_1^* \; E_1)^{-1} & & \\ & \ddots & \\ & & (E_k^* \; E_k)^{-1} \end{pmatrix} E^*$$

with the off diagonal submatrices in the middle matrix being null, from which we obtain

$$E^{-1}E^{*-1} = (E^*E)^{-1} = \begin{pmatrix} (E_1^* \; E_1)^{-1} & & \\ & \ddots & \\ & & (E_k^* \; E_k)^{-1} \end{pmatrix}$$

and taking inverses of both sides,

$$E^*E = \begin{pmatrix} (E_1^* \; E_1) & & \\ & \ddots & \\ & & (E_k^* \; E_k) \end{pmatrix} \quad \text{... (2.9)}$$

The left hand side of (2.9) is

$$\begin{bmatrix} E_1^*E_1 & E_1^*E_2 & \dots & E_1^*E_k \\ \cdot & \cdot & \dots & \cdot \\ E_k^*E_1 & E_k^*E_2 & \dots & E_k^*E_k \end{bmatrix} \qquad \dots \quad (2.10)$$

Comparing (2.9) and (2.10), $E_i^* E_j = 0$, $i \neq j$ which implies $A_i^* A_j = 0$, $i \neq j$.

Corollary: *Let* Λ *be a positive definite matrix and let* A_i *be as defined in Theorem* 2c *with* $\Sigma(\text{rank } A_i) = m$. *Then the two statements*

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

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

are equivalent.

Theorem 2d: *A necessary and sufficient condition for the equation* $AXB = C$ *to have a solution is*

$$AA^-CB^-B = C, \qquad \dots \quad (2.11)$$

in which case the general solution is

$$X = A^-CB^- + Z - A^-AZBB^-. \qquad \dots \quad (2.12)$$

Suppose there exists an X such that $AXB = C$. Then

$$C = AXB = (AA^-A)\,X\,(BB^-B)$$
$$= AA^-(AXB)B^-B$$
$$= AA^-CB^-B$$

establishing the necessity of (2.11). The sufficiency is obvious, for if (2.11) holds, then $X = A^-CB^-$ is a solution.

A^-CB^- is a particular solution of the equation $AXB = C$. To get the general solution we add to it a general solution of the equation $AXB = 0$. It may be een that

$$AXB = 0 \Longleftrightarrow XB = (A^-A - I)FB$$

where F is arbitrary. Further

$$XB = (A^-A - I)\,FB$$
$$\Longleftrightarrow X = (I - A^-A)FBB^- + G(I - BB^-)$$

where G is arbitrary. Now

$$X = FBB^- + G(I - BB^-) - A^-AFBB^-$$
$$= Z - A^-AZBB^-$$

where $Z = FBB^- + G(I - BB^-)$ which is arbitrary since F and G are arbitrary.

Theorem 2e : *A necessary and sufficient condition for the equations* $AX = C$, $XB = D$ *to have a common solution is that the individual equations should have solutions and* $AD = CB$.

Necessity is obvious. To prove sufficiency consider

$$X = A^-C+DB^- -A^-ADB^-$$

which is a solution if the required conditions $AA^-C = C$, $DB^-B = D$ and $AD = CB$ are satisfied.

Theorem 2f : (*Representation of all g-inverses*). *The general solution to a g-inverse of a given* $m\times n$ *matrix* A *is*

$$A^-+U-A^-AUAA^- \quad \dots \quad (2.13)$$

where A^- *is a particular inverse and* U *is an arbitrary* $n\times m$ *matrix.*

The expression (2.13) is the general solution of the equation $AXA = A$, which is deduced from the result of Theorem 2d.

From (d) of Theorem 2a and the result (2.13) of Theorem 2f, it follows that any solution of the consistent equation $AX = Y$ can be written as A^-Y where A^- is a g-inverse of A.

3. A g-inverse for a minimum norm solution of $AX = Y$ (consistent)

In Section 2 we have seen that a general solution of the consistent equation, $AX = Y$, can be expressed in the form

$$X = A^-Y+(I-H)Z \quad \dots \quad (3.1)$$

where A^- is any g-inverse of A, i.e., $AA^-A = A$ as defined in (2.1), and $A^-A = H$. There is no unique solution unless $H = I$ and all solutions are generated by assigning different values to Z. Among these solutions, there exists one which has the smallest norm. Such a solution is called the *minimum norm solution* and can be found by choosing Z such that the norm of (3.1) is a minimum.

We now raise the following question. Does there exist a particular choice of a g-inverse G such that $X = GY$ is a solution with a minimum norm ? Let G exist. Since G is a g-inverse, a general solution of $AX = Y$ is

$$GY+(I-GA)Z \quad \dots \quad (3.2)$$

and by hypothesis

$$\|GY\| \leqslant \|GY+(I-GA)Z\| \quad \text{for all } Z \quad \dots \quad (3.3)$$

which implies that

$$Y^*G^*(I-GA) = 0. \quad \dots \quad (3.4)$$

Since the condition (3.4) should be true for all Y such that $Y = Ab$ where b is arbitrary it follows that

$$A^*G^*(I-GA) = 0 \text{ or } (GA)^*(I-GA) = 0. \quad \dots \quad (3.5)$$

We establish the following lemma.

Lemma 3a : *The necessary and sufficient conditions that $X = GY$ is a minimum norm solution of the consistent equation $AX = Y$ are*

$$\text{(a)} \quad AGA = A, \qquad \text{(b)} \quad (GA)^*(I-GA) = 0. \qquad \dots \quad (3.6)$$

There are, however, different sets of equivalent conditions for a minimum norm solution which we give in Theorem 3a.

Theorem 3a : *The set of conditions*

(i) $AGA = A, \ (GA)^*(I-GA) = 0$

(ii) $AGA = A, (GA)^* = GA,$

(iii) $GAA^* = A^*,$

(iv) $AGA = A, \ GA \epsilon M(A^*),$

are equivalent for GY to be a minimum norm solution of $AX = Y$.

From (i), $(GA)^* = (GA)^*(GA) \implies (GA) = (GA)^*(GA)$ and hence $(GA)^* = (GA)$ which shows that (i) $\implies$ (ii).

From (ii), $A^*G^*A^* = A^* \implies (GA)^*A^* = GAA^* = A^*$, i.e., (ii) $\implies$ (iii).

From (iii), $GAA^*G^* = A^*G^* \implies (GA)^* = GA \implies GA \epsilon M(A^*)$.

Further $AGA = A(GA)^* = AA^*G = A$, which shows that (iii) $\implies$ (iv).

From (iv), $GA = A^*D$ for some matrix D. Then

$$(GA)^* = D^*A \implies (GA)^*(I-GA) = D^*A(I-GA) = 0.$$

Thus (iv) $\implies$ (i) which proves Theorem 3a.

We denote a g-inverse which provides a minimum norm solution of a consistent equation $AX = Y$ by A_m^-. If such an inverse exists then it must satisfy the necessary and sufficient conditions of Lemma 3a.

Corollary : *One choice of A_m^- is $A^*(AA^*)^-$ where $(AA^*)^-$ is any g-inverse of AA^*.*

From (iii) of Theorem 3a it is seen that A_m^- satisfies the equation in G,

$$GAA^* = A^* \qquad \dots \quad (3.7)$$

which is obviously consistent. Hence a solution for G can be obtained as

$$G = A^*(AA^*)^-. \qquad \dots \quad (3.8)$$

This may be independently verified, for by substituting $A^*(AA^*)^-$ for G on the left hand side of (3.7) we obtain

$$A^*(AA^*)^-AA^* \qquad \dots \quad (3.9)$$

which is equal to A^* using (b) of Theorem 2b. The solution (3.8) is not unique unless (AA^*) is of full rank. However for any choice of G, the minimum norm solution of a consistent equation $AX = Y$ is unique in virtue of (vii) of Theorem 2a.

The most general solution of (3.7) is one in which the i-th row of G is $\alpha_i(AA^*)^-$ where α_i is the i-th row of A^* and $(AA^*)^-$ is any g-inverse of (AA^*) and the choice may depend on i. In the solution (3.8), the choice of $(AA^*)^-$ is the same for all i.

The following results are easy to verify for any choice of A_m^-

Theorem 3b :

(i) *One choice of* $(AA^*)_m^-$ *is* $(A_m^-)^*A_m^-$.

(ii) $(AA^*)_m^-AA^* = (AA_m^-)^*$.

(iii) *One choice of* $(UAV)_m^-$ *is* $V^*A_m^-U^*$ *when* U *and* V *are unitary matrices* (i.e., $UU^* = I$, $VV^* = I$).

(iv) *One choice of* $(\lambda A)_m^- = \lambda^{-1}A_m^-$ *where* λ *is a non-zero scalar.*

(v) $(A^*A)_m^-A^*$ *is a g-inverse of* A.

(vi) $A(A^*A)_m^-A^*$ *is symmetric.*

(vii) *If* G_1 *and* G_2 *are any two different choices of* A_m^-, *then* $(G_1-G_2)A = 0$.

(viii) *If rank* $A = n = min(m, n)$ *then every* A^- *is a* A_m^-.

4. A g-inverse for a least square solution

Let us consider an inconsistent equation $AX = Y$. We say that $\hat{X}$ is a least square solution of $AX = Y$ iff

$$\|A\hat{X}-Y\| = \min_X \|AX-Y\| \qquad \dots \quad (4.1)$$

It is well known that $\hat{X}$ exists which, however, may not be unique. Now we ask the question whether there exists a g-inverse, G such that GY is a least square solution of $AX = Y$ for any given Y. Suppose such a matrix G exists. Then

$$\|Y-AGY\| \leqslant \|Y-AGY+AGY-AX\|$$

which implies that the inner product

$$(Y-AGY,\ AGY-AX) = 0 \text{ for all } X \text{ and } Y$$

or

$$A^*(I-AG) = 0.$$

We establish the following lemma.

Lemma 4a : *The necessary and sufficient condition that* $X = GY$ *is a least square solution of the equation* $AX = Y$ *is*

$$A^*(I-AG) = 0. \qquad \dots \quad (4.2)$$

As in the case of a minimum norm solution, there are different sets of equivalent conditions for a least square solution, which are stated in Theorem 4a.

Theorem 4a : *The following sets of conditions*

(i) $A^* = A^*AG$,

(ii) $AGA = A$, $(AG)^* = AG$,

(iii) $AGA = A$, $(AG)^*(I-AG) = 0$,

(iv) $AGA = A$, $(AG)^* \epsilon\, \mathcal{M}(A)$,

are equivalent for GY *to be a least square solution of* $AX = Y$.

The proof of this theorem is similar to that of Theorem 3a and is therefore omitted.

We denote by A_l^-, a g-inverse which provides a least square solution of $AX = Y$, which may be consistent or not. A_l^- is not, however, unique.

Corollary : *One choice of* A_l^- *is* $(A^*A)^-A^*$ *where* $(A^*A)^-$ *is any g-inverse of* (A^*A).

The result follows from condition (i) of Theorem 4a. For further comments see the corollary following Theorem 3a. The following results are easy to verify.

Theorem 4b :

(i) $(\lambda A)_l^- = \lambda^{-1} A_l^-$ *where* λ *is a non-zero scalar.*

(ii) *One choice of* $(UAV)_l^- = V^*A_l^-U^*$ *when* U *and* V *are unitary matrices.*

(iii) *If rank* $A = m = \min(m, n)$, *then every* A^- *is a* A_l^-

(iv) *If* G_1 *and* G_2 *are two choices of* A_l^-, *then* $A(G_1-G_2) = 0$.

5. A g-INVERSE FOR MINIMUM NORM LEAST SQUARE SOLUTION

Let $\hat{X}$ be a least square solution of $AX = Y$. Then

$$\|Y-A\hat{X}\| \leqslant \|Y-AX\| \quad \text{for all } X \qquad \dots \quad (5.1)$$

$$\Longleftrightarrow A^*(Y-A\hat{X}) = 0 \qquad \dots \quad (5.2)$$

so that $\hat{X}$ satisfies the equation

$$A^*AX = A^*Y. \qquad \dots \quad (5.3)$$

The equation (5.2) does not have a unique solution unless the matrix A^*A is nonsingular. However $\|Y-A\hat{X}\|$ is unique for any solution of (5.2). We define X_m to be a minimum norm least square solution if X_m is a least square solution and

$$\|X_m\| \leqslant \|\hat{X}\| \qquad \dots \quad (5.4)$$

for any other least square solution $\hat{X}$.

Let $X = GY$ be a particular solution valid for any given Y. A general solution of (5.3) is

$$GY+(I-(A^*A)^-A^*A)Z \qquad \dots \quad (5.5)$$

where $(A^*A)^-$ is any g-inverse of (A^*A). Since GY is a solution of (5.3)

$$A^*AGY = A^*Y \Longrightarrow A^*AG = A^*. \qquad \dots \quad (5.6)$$

Now, if GY is a minimum norm solution of (5.3), then

$$G^*(I-(A^*A)^-A^*A) = 0$$

$$\Longleftrightarrow G^* = DA, \text{ or } G \in M(A^*), \qquad \dots \quad (5.7)$$

where D is some matrix, using the result of Lemma 2b, since $(A^*A)^-A^*$ is a g-inverse of A. We establish the following lemma.

Lemma 5a : *The necessary and sufficient conditions that GY is a minimum norm least square solution of $AX = Y$ are*

(a) $A^* = A^*AG$, (b) $G \in M(A^*)$ *or* $G^* = DA$ *for some matrix* D. ... (5.8)

Now we consider a number of sets of equivalent conditions which define a g-inverse for the minimum norm least square solution.

Theorem 5a : *The following sets of conditions*

(i) $A^*AG = A^*$, $G^* = DA$ *for some matrix* D,

(ii) $GAA^* = A^*$, $G^* = AD$ *for some matrix* D,

(ii) $GAA^* = A^*$, $AGG^* = G^*$,

(iv) $A^*AG = A^*$, $G^*GA = G^*$,

(v) $AGA = A$, $GAG = G$, $(AG)^* = AG$, $(GA)^* = GA$,

are equivalent for GY to be a minimum norm least square solution of $AX = Y$.

The conditions (iv) are proposed by Moore (1935) and conditions (v) by Penrose (1955).

It is already shown in Theorem 4a that

$$A^*AG = A^* \Longrightarrow AGA = A, \ (AG)^* = AG. \qquad \dots \quad (5.9)$$

Now $\quad DA(I-GA) = DA-DAGA = 0 \Longrightarrow G^*(I-GA) = 0.$

Further $\quad G^*(I-GA) = 0 \Longrightarrow A^*G^*(I-GA) = 0$ or

$$(GA)^* = (GA)^*GA = GA. \qquad \dots \quad (5.10)$$

Also, $G^*(I-GA) = 0 \Longrightarrow G^*(I-A^*G^*) = 0$ using (5.10) $\Longrightarrow G^* = G^*A^*G^*$ or

$$GAG = G. \qquad \dots \quad (5.11)$$

The results (5.9), (5.10) and (5.11) show that (i) $\Longrightarrow$ (v). From (v),

$$A^*G^*A^* = A^* \Longrightarrow GAA^* = A^* \text{ using } A^*G^* = GA$$

$$G^* = G^*A^*G^* = AGG^* = AD, \text{ using } G^*A^* = AG.$$

Hence (v) $\Longrightarrow$ (ii). From (ii)

$$GAA^* = A^* \Longrightarrow AGA = A.$$

Then

$$G^* = AD = AG\,AD = AGG^*$$

which shows that (ii) $\Longrightarrow$ (iii).

From (iii), $GAA^* = A^* \Longrightarrow AGA = A$. Now

$$A^*AG - A^*)(G^*A^*A - A) = 0$$

using the conditions $AGA = A$ and $AGG^* = G^*$. Hence $A^*AG = A^*$. Similarly $G^*GA = G^*$. Thus (iii) $\Longrightarrow$ (iv). It is easy to see that (iv) $\Longrightarrow$ (i). Thus Theorem 5a is established.

A g-inverse satisfying any of the sets of conditions in Therem 5a is called Moore-Penrose inverse and is represented by A^+ using the notation used by Penrose. According to the notation of this paper the symbol A^-_{lm} will be used for A^+.

Lemma 5b : *A^+ is unique.*

Let G be any matrix satisfying the conditions (5.8). Then using the condition (a) of (5.8)

$$A^*A(G - A^+) = 0 \text{ or } A(G - A^+) = 0. \qquad \dots (5.12)$$

Using condition (b) of (5.8)

$$(G - A^+)^* = (D_1 - D_2)A$$
$$\Longrightarrow T(G - A^+) = 0 \qquad \dots (5.13)$$

where the rows of T are orthogonal to the rows of A. The equations (5.12) and (5.13) together imply that $G - A^+ = 0$ which proves the desired result.

The following results are easy to verify.

Theorem 5b :

(i) $(A^+)^+ = A$.

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

(iii) $(AA^*)^+ = (A^+)^*A^+$.

(iv) *rank* $A =$ *rank* A^+.

(v) *rank* $A^+ =$ *rank* A^+A.

(vi) $A^+A = AA^+$ *if A is normal.*

(vii) $(A^n)^+ = (A^+)^n$ *if A is normal.*

(viii) $(AA^*)^+(AA^*) = (AA^+)^*$.

(ix) $(UAV)^+ = V^*A^+U^*$ *when U and V are unitary matrices.*

(x) $(\lambda A)^+ = \lambda^{-1}A^+$.

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

(xii) $AA^+ = AA^-_l$ *for any choice of* A^-_l.

(xiii) $A^+A = A^-_mA$ *for any choice of* A^-_m.

(xiv) $A^+ = (A^*A)^-_mA^*$ *for any choice of* $(A^*A)^-_m$.

(xv) $A^+ = A^*(AA^*)^-_l$ *for any choice of* $(AA^*)^-_l$.

(xvi) $A^+ = A^-_m$ *if rank* $A = m = \min(m, n)$.

(xvii) $A^+ = A^-_l$ *if rank* $A = n = \min(m, n)$.

(xviii) *Let A and B be any two matrics with the product AB defined.* Furthermore *let $B_1 = A^+AB$ and $A_1 = AB_1B_1^+$. Then $AB = A_1B_1$ and* $(AB)^+ = B_1^+A_1^+$ (Cline, 1964).

6. PROJECTION OPERATOR

Let B be a $(m \times m)$ matrix which projects (orthogonally) vectors onto the manifold $M(A)$, generated by the columns of a $m \times n$ matrix A. Choose an arbitrary vector Y. Then

$$BY \epsilon M(A) \text{ for all } Y \Longrightarrow B = AD \qquad \ldots \quad (6.1)$$

where D is $(n \times m)$ matrix. Further $(Y-BY)$ is orthogonal to $M(A)$ for any Y, i.e.,

$$A^*(I-B) = 0. \qquad \ldots \quad (6.2)$$

We establish the following theorem.

Theorem 6a: *The necessary and sufficient conditions that B projects vectors onto $M(A)$ are*

$$\text{(a)} \quad B = AD, \qquad \text{(b)} \quad A^* = A^*B \qquad \ldots \quad (6.3)$$

which are equivalent to the conditions

$$\text{(a)} \quad (AD)^* = AD, \qquad \text{(b)} \quad ADA = A \qquad \ldots \quad (6.4)$$

where D is as defined in (6.1).

From (a) and (b) of (6.3), $B^* = D^*A^* = D^*A^*B = B^*B$. Hence $B = AD$ is symmetrical. Further

$$A^* = A^*B = A^*B^* = A^*D^*A^*$$

which proves that the conditions of (6.3) imply the conditions of (6.4). The reverse is easy to prove.

Corollary 1: *A projection operator is symmetric and idempotent.*

Corollary 2: *An operator projecting vectors onto $M(A)$ can be written as AA_l^-.*

From the conditions (6.4) we see that D is A_l^-. Hence $B = AA_l^-$.

Corollary 3: *$A(A^*A)^-A^*$ is symmetric, idempotent and projects vectors onto $M(A)$.*

The result follows since one choice of A_l^- is $(A^*A)^-A^*$.

7. g-INVERSE FOR A BASIC SOLUTION OF $AX = Y$ (CONSISTENT)

We say that X_b is a basic solution of $AX = Y$ if

(i) $AX_b = Y$, and

(ii) X_b has at most r non-zero components where r is rank A.

Let G be an inverse which provides X_b. Then GY has $(n-r)$ components zero, whenever $Y \epsilon M(A)$, i.e., GA has $(n-r)$ null rows. Let us suppose without loss of generality that the last $n-r$ rows are null. Then writing G as a partitioned matrix we have

$$\begin{pmatrix} G_1 \\ \cdots \\ G_2 \end{pmatrix} A = \begin{pmatrix} G_1 A \\ 0 \end{pmatrix} \qquad \ldots \quad (7.1)$$

where G_1 has r rows, giving

$$G_2 A = 0. \qquad \ldots \quad (7.2)$$

Further writing A as a partitioned matrix the equation $AGA = A$ becomes

$$\left(A_1 \vdots A_2\right) \begin{pmatrix} G_1 \\ \cdots \\ G_2 \end{pmatrix} A = A \qquad \ldots \quad (7.3)$$

giving

$$A_1 G_1 A = A \qquad \ldots \quad (7.4)$$

where A_1 has r columns, which implies that rank $G_1 =$ rank $A_1 = r$ and $A_2 \epsilon M(A_1)$. The equation (7.4) can be written as

$$A_1 G_1 A_1 = A_1, \qquad A_1 G_1 A_2 = A_2 \qquad \ldots \quad (7,5)$$

which shows that G_1 is a $(A_1)^-$, the second equation being automatically satisfied since $A_2 \epsilon M(A_1)$.

Multiplying the first equation of (7.5) by A_1^*

$$A_1^* A_1 G_1 A_1 = A_1^* A_1 \Longrightarrow G_1 A_1 = I \qquad \ldots \quad (7.6)$$

since $A_1^* A_1$ is a square matrix with full rank. Hence G_1 is also $(A_1)_m^-$. Thus we establish Lemma 7a.

Lemma 7a : *The necessary and sufficient condition that G is a g-inverse of A providing a basic solution of a consistent equation $AX = Y$ is*

$$G_1 = (A_1)^-, \; G_2 A = 0 \qquad \ldots \quad (7.7)$$

where G_1, G_2 are partitions of G with r and $n-r$ rows respectively, A_1 is $m \times r$ matrix formed by any r independent columns of A (which may be taken as the first r columns of A by suitable rearrangement), $(A_1)^-$ is any g-inverse of A_1 and G_2 is any matrix such that $G_2 A = 0$. In fact $(A_1)^- = (A_1)_m^-$.

We represent such an inverse by A_b^-. It is seen that if we want a solution with at most $s > r$ nonzero components, then the corresponding g-inverse can be written as

$$\begin{pmatrix} (A_1)^- \\ \\ G_2 \end{pmatrix} \qquad \ldots \quad (7.8)$$

where $(A_1 \vdots A_2)$ is a partition of A such that A_1 contains s columns out of which r are independent and G_2 is any matrix such that $G_2 A = 0$. In such a case $(A_1)^-$ is no longer a $(A_1)_m^-$.

8. A g-inverse for a basic least square solution of $AX = Y$

A vector X_b is said to be a basic least square solution of $AX = Y$ if

(i) $\|AX_b - Y\| \leqslant \|AX - Y\|$ for all X, and

(ii) X_b has at most r nonzero components, where r is the rank of A.

Let G be an inverse of A which provides such a solution. Then G satisfies the conditions of A_l^-. In addition GY can have at most r nonzero components for any Y, which implies that $(n-r)$ rows of G are null. Let us suppose without loss of generality that the last $(n-r)$ rows of G are null in which case G can be written as

$$\begin{pmatrix} G_1 \\ \cdots \\ 0 \end{pmatrix}$$

where G_1 has r rows. Let us consider the corresponding partition of $A = (A_1 \vdots A_2)$ where A_1 has r columns and A_2 has $(n-r)$ columns. Then we have the following lemma concerning the representation of a g-inverse which provides a basic least square solution and which may be denoted by A_{lb}^-

Lemma 8a : *A_{lb}^- can be expressed in the form*

$$\begin{pmatrix} (A_1)_l^- \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} A_1^\dagger \\ \cdots \\ 0 \end{pmatrix} \qquad \ldots \quad (8.1)$$

where A_1 is any set of r independent columns of A which may be written as the first partition of A by interchanging columns if necessary.

Using the partitioned forms of A and G, we have

$$\left(A_1 \vdots A_2\right) \begin{pmatrix} G_1 \\ \cdots \\ 0 \end{pmatrix} A = A \quad \text{or} \quad A_1G_1A = A \qquad \ldots \quad (8.2)$$

which shows that rank A_1 = rank G_1 = rank $A = r$, since A_1 has r columns and G_1 has r rows. The equation (8.2) can be written

$$A_1G_1A_1 = A_1,\ A_1G_1A_2 = A_2. \qquad \ldots \quad (8.3)$$

The first equation shows that G_1 is a g-inverse of A_1 and the second equation is automatically satisfied since $A_2 \epsilon M(A_1)$. Further

$$AG = A_1G_1 \Longrightarrow (A_1G_1)^* = (A_1G_1) \qquad \ldots \quad (8.4)$$

since G satisfies the conditions of A_l^-. Hence G_1 is $(A_1)_l^-$. Since rank A_1 is equal to the number of columns of A_1, $(A_1)_l^- = A_1^\dagger$ which proves the lemma.

It may be noted that A_{lb}^- is not unique, since A_1 may be chosen in different ways.

Lemma 8b : *A_{lb}^- satisfies the reflexive condition in addition to being A_{lb}^-.*

We have only to verify that the reflexive condition is satisfied by A_{lb}^-, i.e.,

$$A_{lb}^- AA_{lb}^- = A_{lb}^-. \qquad \ldots \quad (8.5)$$

Using the representation (8.1) of A_{lb}^- as found in Lemma 7a

$$\begin{pmatrix} A_1^\dagger \\ 0 \end{pmatrix} (A_1 \vdots A_2) \begin{pmatrix} A_1^\dagger \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} A_1^\dagger A_1 & A_1^\dagger A_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_1^\dagger \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} A_1^\dagger A_1 A_1^\dagger \\ 0 \end{pmatrix} = \begin{pmatrix} A_1^\dagger \\ 0 \end{pmatrix}$$

which establishes the equation (7.5).

The pseudoinverse constructed by Rao (1955) is of the type A_{lb}^-. Lancoz (1961) has considered such inverses and Rosen (1964) provided an algorithm for computing them.

9. A g-INVERSE HAVING EIGEN VALUES RECIPROCAL TO THOSE OF A

When A is a square nonsingular matrix the eigen values of A^{-1} are reciprocals of those of A and the eigen vectors of A are also the eigen vectors of A^{-1}. Is it possible to construct a g-inverse, when A is a square singular matrix, for which a similar property holds ? It is clear that if there exists a g-inverse which commutes with A, then the above property is satisfied. We state another sufficient condition in Lemma 9a.

Lemma 9a : *Let G be a generalized inverse of a square matrix A,* i.e., $AGA = A$. *If*

$$G \sum_{i=1}^{k} a_i A^i = \sum_{i=1}^{k} a_i A^{i-1}$$

for some k and constants $a_1, \ldots, a_k$, then

$$AX = \lambda X \Longrightarrow GX = \lambda^{-1} X$$

provided $\Sigma a_i \lambda^i \neq 0$.

The result immediately follows since

$$G \Sigma a_i A^i X = \Sigma a_i A^{i-1} X$$

$$G \Sigma a_i \lambda^i X = \Sigma a_i \lambda^{i-1} X$$

or dividing by $\Sigma a_i \lambda^i$,

$$GX = \lambda^{-1} X.$$

We have given only sufficient conditions for a g-inverse to have eigen values which are reciprocals of the eigen values of A. An inverse with this property is denoted by A_e^-. Methods of constructing such inverses are given by Drazin (1958) and Scroggs and Odell (1966).

10. g-INVERSE WITH MAXIMUM RANK

If a general inverse G of A satisfies the conditions $AGA = A$ and $GAG = G$, then it follows that rank A = rank G. However, if $AGA = A$ is the only condition to be satisfied, then there is a possibility of rank G exceeding rank A. We show that there exists a g-inverse with the maximum rank which is min (m, n) whatever may be rank A.

Let $m \leqslant n$ and consider the square matrix B obtained by adding $(n-m)$ zero rows to A. It is known that there exists a $n \times n$ non-singular matrix C such that $CB = H$, where H is in Hermite canonical form with the property $H^2 = H$ (Rao, 1965, p. 18).

Consider the $(n \times m)$ matrix A^- obtained by omitting the last $(n-m)$ columns of C. Obviously rank A^- is m and we show that A^- is a g-inverse. Now

$$CBCB = H^2 = H$$

$$\Longrightarrow BCB = C^{-1}H = B = \begin{pmatrix} A \\ 0 \end{pmatrix} \qquad \ldots \quad (10.1)$$

$$BCB = \begin{pmatrix} A \\ 0 \end{pmatrix} (A^- \vdots D) \begin{pmatrix} A \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} AA^- & AD \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A \\ 0 \end{pmatrix} = \begin{pmatrix} AA^-A \\ 0 \end{pmatrix} \qquad \ldots \quad (10.2)$$

where C is written $(A^- \vdots D)$. Comparing (10.1) and (10.2) $AA^-A = A$, i.e., A^- is a g-inverse.

If $n<m$, B is obtained by adding $(m-n)$ zero columns to A and A^- by omitting the last $(m-n)$ rows of C.

A g-inverse *so constructed* with the maximum rank provides a test for consistency of given linear equations different from the one proposed in Theorem 2a based on any g-inverse.

Lemma 10a : *Consider A^- as defined in* (10.1, 10.2). *Let $A^-A = H$ which is $n \times n$ square matrix, and $A^-Y = h$ which is $n \times 1$ column vector. The necessary and sufficient condition that $AX = Y$ is consistent is that the i-th component of h is zero if the i-th row of H consists of all zeroes.*

It may be seen that under the condition of the Lemma 9a rank A = rank $(A \vdots Y)$ which is the condition for the existence of a solution of $AX = Y$. A g-inverse with maximum rank as defined in (10.1, 10.2) is denoted by A_c^-.

11. Reflexive type g-inverse

In addition to the g-inverses defined in the earlier sections we shall consider a few more which are not related to any specific purpose but which are interesting. It may be noted that the reflexive condition $GAG = G$ did not play any part in many of the definitions. We can, however, introduce it as an additional condition in some situations. For instance, we can define a g-inverse by the conditions $AGA = A$, $GAG = G$. Such an inverse may be denoted by A_r^- (a g-inverse which is reflexive). Similarly by introducing the additional condition $GAG = G$, we define A_{mr}^-, A_{lr}^-. The following lemmas contain sets of equivalent conditions for such inverses.

Lemma 11a : *The following sets of conditions*

(i) $AGA = A$, $G = A^*D$ *for some matrix* D,

(ii) $AGA = A$, $GAG = G$, $(GA)^* = GA$,

are equivalent.

The inverse so obtained is an A_m^- with the additional reflexive condition $GAG = G$ and is denoted by A_{mr}^-. The equivalence of the conditions is easily established.

Lemma 11b : *The following sets of conditions*

(i) $AGA = A$, $G = DA^*$ *for some matrix* D,

(ii) $AGA = A$, $GAG = G$, $(AG)^* = AG$,

are equivalent.

The inverse so obtained is an A_l^- with the additional reflexive condition $GAG = G$ and is denoted by A_{lr}^-. Again, the equivalence of the conditions is easily established. The definition (ii) of Lemma 11b was proposed by Zelen and Goldman (1964) under the name of 'weak generalized inverse'.

Lemma 11c : *A general representation of* A_{lr}^- *is* DA^* *where*

$$D = D_0+Z-D_0A^*AZA^*AD_0$$

Z being an arbitrary matrix and $D_0 = (A^*A)^-$, *any g-inverse of* (A^*A).

Let $A_{lr}^- = DA^* = (D_1+D_0)A^*$. Then using the condition $AA_{lr}^- A = A$

$$AD_1A^*A = 0 \Longleftrightarrow A^*AD_1A^*A = 0.$$

A general solution of D is then, using the result of Theorem 2d,

$$D = D_0+Z-D_0A^*AZA^*AD_0$$

where Z is arbitrary. Thus A_{lr}^- is not unique and the least square solution based on such an inverse is also not unique.

12. A CLASSIFICATION OF DIFFERENT TYPES OF g-INVERSES

Let G denote a g-inverse of A. The different types of g-inverses studied in Sections 2-11 and their definitions are brought together in this section and exhibited in a tabular form.

TABLE 1. NON-REFLEXIVE TYPE g-INVERSES

purpose	sets of equivalent conditions	symbol
1. solutions of $AX=Y$, $AXB=C$, etc.	$AGA=A$	A^-
2. minimum norm solution of $AX=Y$ (consistent)	(i) $AGA=A$, $(GA)^*(I-GA)=0$ (ii) $AGA=A$, $(GA)^*=GA$ (iii) $GAA^*=A^*$ (iv) $AGA=A$, $GA\epsilon M(A^*)$	A_m^-
3. least square solution of $AX=Y$ (inconsistent)	(i) $A^*=A^*AG$ (ii) $AGA=A$, $(AG)^*=AG$ (iii) $AGA=A$, $(AG)^*(I-AG)=0$ (iv) $AGA=A$, $(AG)^*\epsilon M(A)$	A_l^-
4. basic solution of $AX=Y$ (consistent)	$G=\begin{pmatrix} G_1 \\ \cdots \\ G_2 \end{pmatrix}$, $A=(A_1 : A_2)$ rank G_1 = rank A_1 = rank G $G_1=(A_1)^-$, $G_2A=0$	A_b^-
5. inverse with the same eigen vectors and eigen values reciprocal to those of A	$AGA=A$ $G\Sigma a_iA^i=\Sigma a_iA^{i-1}$	A_e^-
6. inverse with the maximum rank for testing consistency of equations	$AGA=A$ rank $G=\min(m, n)$	A_c^-

TABLE 2. REFLEXIVE TYPE g-INVERSES

	purpose	sets of equivalent conditions	symbol
1.	solutions of $AX=Y$, $AXB=C$, etc.	$AGA=A$, $GAG=G$	A_r^-
2.	minimum norm solution of $AX=Y$ (consistent)	(i) $AGA=A$, $G\epsilon \mathcal{M}(A^*)$ (ii) $AGA=A$, $GAG=G$, $(GA)^*=GA$	A_{mr}^-
3.	least square solution of $AX=Y$ (inconsistent)	(i) $AGA=A$, $G^*\epsilon \mathcal{M}(A)$ (ii) $AGA=A$, $GAG=G$, $(AG)^*=AG$	A_{lr}^-
4.	basic least square solution of $AX=Y$ (inconsistent)	$G=\begin{pmatrix} G_1 \\ \cdots \\ G_2 \end{pmatrix}$, $A=(A_1 \vdots A_2)$ rank $G_1=$ rank $A_1=$ rank A $G_1=(A_1)^+=(A_1)_l^-$, $G_2=0$	A_{lb}^-
5.	minimum norm least square solution of $AX=Y$ (inconsistent)	(i) $A^*AG=A^*$, $G\epsilon \mathcal{M}(A^*)$ (ii) $GAA^*=A^*$, $G^*\epsilon \mathcal{M}(A)$ (iii) $GAA^*=A^*$, $AGG^*=G^*$ (iv) $A^*AG=A^*$, $G^*GA=G^*$ (v) $AGA=A$, $GAG=G$, $(AG)^*=AG$, $(GA)^*=GA$.	A^+

13. Some further examples of g-inverses

(i) Let A be $m\times n$ matrix of rank r. Consider the partitions

$$A=\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}=(B_1 \vdots B_2)=\begin{pmatrix} C_1 \\ \cdots \\ C_2 \end{pmatrix}$$

such that A_1 is $r\times r$ matrix of rank r, B_1 is $m\times r$ matrix of rank r and C_1 is $r\times n$ matrix of rank r, by suitable rearrangement of columns and rows if necessary. It may be seen that

$$A_4=A_3A_1^{-1}A_2.$$

It is easy to verify the following

(a) $\begin{pmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ is A_b^- and reflexive.

(b) $\begin{pmatrix} A_1^{-1} & -K \\ 0 & I \end{pmatrix}$ is A_e^-, where K is such that $A_1K=A_2$.

(c) $\begin{pmatrix} B_1^- \\ 0 \end{pmatrix}$ is A_b^- and reflexive.

(d) $(C_1^- \vdots 0)$ is A_r^-.

(e) $\begin{pmatrix} (B_1)_l^- \\ 0 \end{pmatrix}=\begin{pmatrix} B_l^+ \\ 0 \end{pmatrix}$ is A_{lb}^- and reflexive.

(f) $((C_1)_m^- \vdots 0) = (C_1^+ \vdots 0)$ is A_{mr}^-.

(g) $D^*(DD^*)^{-1}B_1^+$ is A^+ where $D = B_1^+A$ and $B_1^+ = (B_1)_l^-$.

(h) $C_1^+(D^*D)^{-1}D^*$ is A^+ where $D = AC_1^+$, $C_1^+ = C_1^*(C_1C_1^*)^{-1}$.

(i) $A^*(AA^*)^-$ is A_{mr}^-.

(j) $(A^*A)^-A^*$ is A_{lr}^-.

(k) $A^*(AA^*)^{-1}$ is A^+ if rank $A = m < n$.

(l) $(A^*A)^{-1}A^*$ is A^+ if rank $A = n < m$.

(ii) Given a $(m \times n)$ matrix A, there exists an orthogonal Householder matrix Q (see page 20 of Rao, 1965) of order m such that

$$QA = \begin{pmatrix} T \\ \cdots \\ 0 \end{pmatrix}, \quad \text{if } m \geqslant n$$

where T is an upper diagonal matrix of order $(n \times n)$ and 0 is $(m-n \times n)$ matrix of zeroes. If $n < m$, then

$$AQ = (T \vdots 0)$$

where T is lower triangular of order $(m \times m)$. It is easy to verify the following.

If $m \geqslant n$, then

(a) $(T^- \vdots 0)\, Q$ is A^-, and $(T_r^- \vdots 0)Q$ is A_r^-.

(b) $(T^+ \vdots 0)Q$ is A^+ and $(T^{-1} \vdots 0)Q$ is A^+ when T has full rank.

(c) $(T_l^- \vdots 0)Q$ is A_l^-, and $(T_{lr}^- \vdots 0)Q$ is A_{lr}^-.

(d) $(T_m^- \vdots 0)Q = A_m^-$, and $(T_{mr}^- \vdots 0)Q$ is A_{mr}^-.

(e) $(T_{lb}^- \vdots 0)Q = A_{lb}^-$.

If $n \leqslant m$, then

(f) $Q\begin{pmatrix} T^- \\ \cdots \\ 0 \end{pmatrix}$ is A^-, and $Q\begin{pmatrix} T_r^- \\ \cdots \\ 0 \end{pmatrix}$ is A_r^-.

(g) $Q\begin{pmatrix} T^+ \\ \cdots \\ 0 \end{pmatrix}$ is A^+, and $Q\begin{pmatrix} T^{-1} \\ \cdots \\ 0 \end{pmatrix} = A^+$ when T has full rank.

(h) $Q\begin{pmatrix} T_l^- \\ \cdots \\ 0 \end{pmatrix}$ is A_l^-, and $Q\begin{pmatrix} T_{lr}^- \\ \cdots \\ 0 \end{pmatrix}$ is A_{lr}^-.

(i) $Q\begin{pmatrix} T_m^- \\ \cdots \\ 0 \end{pmatrix}$ is A_m^-, and $Q\begin{pmatrix} T_{mn}^- \\ \cdots \\ 0 \end{pmatrix}$ is A_{mr}^-.

(j) $Q\begin{pmatrix} T_{lb}^- \\ \cdots \\ 0 \end{pmatrix}$ is A_{lb}^-.

(iii) Givien a $(m \times n)$ matrix A, there exist nonsingular matrices B and C such that

$$BAC = \Delta$$

where Δ is a diagonal matrix, not necessarily square, which may have nonzero elements only in the main diagonal. Define by Δ^-, the $n \times m$ matrix obtained by considering the transpose of Δ and replacing the nonzero elements by their reciprocals. Then $C\Delta^- B$ is A_r^-.

(iv) In the decomposition $BAC = \Delta$, we can choose B and C to be orthogonal matrices. Then $C\Delta^-B$ is A^+.

(v) Given a square matrix A of order m, there exists a nonsingular matrix T such that

$$TAT^{-1} = J$$

where J is in Jordan canonical form

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

where J_i, the i-th Jordan block, is of the form

$$\begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

where λ_i is an eigen value of A. We define

$$J_i^- = J_i^{-1} \text{ when } \lambda_i \neq 0$$
$$= J_i' \text{ when } \lambda_i = 0$$

and

$$J^- = \begin{bmatrix} J_1^- & & & \\ & J_2^- & & \\ & & \ddots & \\ & & & J_k^- \end{bmatrix}$$

It may be verified that $A^- = TJ^-T^{-1}$ is A_r^- and reflexive.

(v) Consider a matrix of the form A^*A which is non-negative definite and occurs as the matrix of normal equations in the least square theory.

(a) Suppose that it is possible to meet the deficiency in the $(m \times n)$ matrix A by adding the rows of a matrix B of minimum rank such that the rank of the extended matrix is n. Consider

$$(A^* \vdots B^*)\begin{pmatrix} A \\ \cdots \\ B \end{pmatrix} = A^*A + B^*B$$

which is of full rank n and which admits a regular inverse. Then

$$(A^*A+B^*B)^{-1} \text{ is } (A^*A)_c^-.$$

(b) Find B as in (a) above with the restriction that rank B is equal to the number of its rows. Observe that the rank of

$$\begin{pmatrix} A^*A & B^* \\ B & 0 \end{pmatrix}$$

is full equal to $2m-r$ where r is the rank of A. Consider

$$\begin{pmatrix} A^*A & B^* \\ B & 0 \end{pmatrix}^{-1} = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}.$$

Then C_1 is a g-inverse of (A^*A).

(c) Suppose that it is possible to write

$$A^*A = \begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix}$$

where S_1 is a submatrix of full rank r. Then

$$\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} \text{ is } (A^*A)_r^-.$$

(d) There exists a nonsingular matrix C such that

$$C(A^*A)C^* = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

by a rearrangement of the columns of A^*A, if necessary. Such a matrix C may be found by the square root method as an upper triangular matrix. Then

$$C^*C = (A^*A)_c^-.$$

Now let C_1 be the matrix obtained by omitting the last $(n-r)$ columns and rows of C. Consider

$$B = \begin{pmatrix} C_1^* & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} C_1^*C & 0 \\ 0 & 0 \end{pmatrix}$$

Then B is $(A^*A)_b^-$ and BA is A_{lb}^-.

(vi) Let

$$M = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}$$

be a partitioned form of a non-negative Hermitian matrix, where A and B are square matrices. The matrix M can be written

$$M = \begin{pmatrix} X_1^* \\ \cdots \\ X_2^* \end{pmatrix} (X_1 \vdots X_2)$$

so that $$A = X_1^*X_1,\ B = X_2^*X_2,\quad C = X_1^*\ X_2.$$

Consider the matrix (Rhodes, 1965)

$$G = \begin{pmatrix} A^-+A^-CD^-C^*A^- & -A^-CD^- \\ -D^-C^*A^- & D^- \end{pmatrix}$$

where $D = B-C^*A^-C$. Then

(a) G is M^-.

(b) G is M_r^- if A^- and D^- are replaced by A_r^- and D_r^- respectively.

(c) G is M^+ if A^- is replaced by A^+ and D is nonsingular.

REFERENCES

(Not all references are cited in the text)

AITKEN, A. C. (1934): On least squares and linear combinations of observations. *Proc. R. Soc. Edin.*, 55, 42-48.

BEN-ISRAEL, A. and CHARNES, A. (1963): Contributions to the theory of generalized inverses. *J. Soc. Ind. Appl. Math.*, 11, 667-699.

BEN-ISRAEL, A. and WERSAN, S. J. (1963): An elimination method for computing the generalized inverse of an arbitrary matrix. *J. Assoc. Comput. Mach.*, 10, 532-537.

BJERHAMMER, A. (1951): Rectangular reciprocal matrices, with special reference to geodetic calculations. *Bull. Geod. Int.*, 52, 188-220.

BLATTNER, J. W. (1963): Bordered matrices. *J. SIAM*, 10, 528-536.

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

BOSE, R. C. (1959): *Analysis of Variance.* Unpublished lecture notes, University of North Carolina, Chapel Hill.

BOTT, R. and DUFFIN, R. J. (1953): On the algebra of net works. *Trans. Am. Math. Soc.*, 74, 99-109.

BUSINGER, P. and GOLUB, G. H. (1965): Linear least square solutions by Householder transformations. *Numerische Mathematik*, 7, 269-276.

CHIPMAN, J. S. and RAO, M. M. (1964): Projections, generalized inverses and quadratic forms. *J. Math. Analysis Applic.*, 9, 1-11.

CLINE, R. E. (1964): Note on the generalized inverse of the product of matrices. *SIAM Review*, 6, 56-58.

——— (1963): Representation of the generalized inverse of matrices with applications in linear programming. Doctoral Thesis, Purdue University.

——— (1964): Representations for the generalized inverse of a partitioned matrix. *J. SIAM*, 12, 588-600.

DAVID, F. N. and NEYMAN, J. (1938): Extension of the Markoff theorem on least squares. *Stat. Res. Mem.*, 2, 105-116.

DESOER, C. A. and WHALEN, B. H. (1963): A note on pseudoinverses. *J. SIAM.*, 11, 442-447.

DRAZIN, M. P. (1958): Pseudoinverses in associative rings and semigroups. *Amer. Math. Monthly*, 65, 506-514.

GREVILLE, T. N. E. (1959): The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations. *SIAM Review*, January, 38-43.

——— (1960): Some applications of the pseudoinverse of a matrix. *SIAM Review*, 2, 15-22.

HESTENES, M. R. (1958): Inversion of matrices by biorthogonalisation and related results. *J. SIAM*, 6, 51-90.

HSU, P. U. (1946): On a factorization of pseudo-orthogonal matrices. Quart. *J. Math.*, Oxford Series, 17, 162-65.

IJIRI, YUJI (1965): On the generalized inverse of an incidence matrix. *J. SIAM*, 13, 827-836.

LANEZOS, C. (1961): *Linear Differential Operators*, Van Nostrand, Princeton, 124-129.

MOORE, E. H. (1935): General analysis. *American Phil. Soc.*, Philadelphia.

MURRAY, F. J. and VON NEUMANN, J. (1936): On rings of operators. *Ann. Math.*, 37, 116-229 (Vol. III, No. 2).

PENROSE, R. (1955): A generalized inverse for matrices. *Proc. Camb. Phil. Soc.*, 51, 406-13.

——— (1956): On best approximate solutions of linear matrix equations. *Proc. Camb. Phil. Soc.*, 52, 17-19.

PYLE, L. D. (1960): The generalized inverse in linear programming. Doctoral Thesis, Purdue University.

RAO, C. RADHAKRISHNA (1955): Analysis of dispersion for multiple classified data with unequal numbers in cells. *Sankhyā*, 15, 253-280.

——— (1961): A study of large sample test criteria through properties of efficient estimates. *Sankhyā*, 23, 25-40.

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

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

——— (1966): Generalized inverse for matrices and its applications in mathematical statistics. *Research Papers in Statistics*, Festschrift for J. Neyman, John Wiley and Sons, New York.

RADO, R. (1956) : Note on generalized inverses of matrices. *Proc. Cam. Phil. Soc.*, 52, 600-601.

RHODES, C. A. (1964) : Contributions to the theory, computation, and application of generalized inverses. Ph.D. thesis, University of North Carolina at Raleigh.

——— (1965) : Generalized inverses of partitioned matrices. *J. SIAM*, 13, 1033-1035.

ROSEN, J. B. (1964) : Minimum and basic solutions to singular linear systems. *J. SIAM*, 12, 156-162.

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

WILKINSON, G. N. (1958) : Estimation of missing values for the analysis of incomplete data. *Biometrics*, 14, 257-286.

ZELEN, M. and GOLDMAN, A. J. (1964) : Weak generalized inverses and minimum variance linear unbiased estimation (unpublished).

Paper received : April, 1967.

BOOK REVIEWS

Proceedings of the Symposium on Congestion Theory : Edited by Walter L. Smith and William E. Wilkinson, University of North Carolina Monograph Series in Probability and Statistics Number 2.

This monograph incorporates 14 papers presented at the Symposium on Congestion Theory held at the University of North Carolina in August 1964. Papers have discussed various aspects of queueing problems that arise in various situations, and each paper has been followed by discussions which form a very important part of the Proceedings. In each paper an attempt has been made to survey the present state of knowledge and also the perspective for future work.

The reviewer would classify the papers under three main groups :(i) papers dealing with new mathematical methods to solve the problems of queueing and congestion; (ii) papers dealing with new angles of queueing problems; (iii) papers with a statisical and practical bias. The paper No. 7 "Markovian Queues" by R. Syski does not really come under any of the three above groups, as it gives a broad survey of the problems connected with Markovian queues and the possible researches in the field.

The papers 1 (by Pollaczek), 2 (by Keilson), 12 (by Takács) and 13 (by Runnenburg) come under group (i). In paper 1 Pollaczek introduces an analytical method of determining the d.f. of waiting time in a multi-channel telephone trunk problem where the d.f. of the service times are known. His method has application in general congestion problems, and perhaps has been presented here for the first time in English language. Runnenburg's method of collective marks (paper 13) gives a lucid way of deducing some known results in queueing theory in an alternative manner, as pointed out by Professor Takacs in discussion; Runnenburg's method can be applied to deduce Pollaczek's equations in a probabilistic way. Takacs (paper 12) gives a general method of dealing with the queueing problems by the application of ballot theorems. A limitation of Takács' method is that it cannot be applied on a random walk with two impenetrable walks; and hence on a dam with finite capacity or a queueing system with limited waiting space or with a bounded waiting time. There is, however, a great scope of applying combinatorial methods for all types of queueing or dam problems.

Under group (ii) come the papers 4 (Saaty), 5 (Heathcote), 6 (Kingman), 8 (Gaver Jr), 9 (Weiss), 11 (Prabhu), 14 (Reich). The papers by Saaty and Reich are related, and deal with an important and unexplored branch of queueing theory, that of network queues (queues in series, parallel, etc.). A study of departure processes is necessary for knowing the behaviour of network queues. A limitation of most of the studies is that whenever the arrival or service process is non-Poisson, it is difficult to derive the d.f. of the waiting time in second or third queue in series. Saaty has shown that there is scope for applying graph theory in net-work queues; this branch is worth exploring.

Heathcote's paper is on divergent single-server queues, which occur when traffic intensity (ρ) is greater than unity. The paper is somewhat long-winded, and the significant properties in this situation could be stated straightway. For example, if $\rho \geqslant 1$, a stable distribution of the waiting time does not exist if there is no restriction in a customer's waiting time or if the waiting space is unlimited; otheriso it is always possible to obtain a stable d.f. for waiting time or queue size. A study of dual queues gives an insight into Heathcote's problems.

Kingman presents the problem of heavy traffic, i.e., the situation when ρ is nearly 1, in a lucid manner. Heavy traffic appears in everyday congestion problems, and his investigations should have a great practical value also.

Gaver presents the problem of priority queues which really form a multi-dimensional stochastic process.

Prabhu brings out the similarities in the models of dam problems and queueing problems, particularly for time-dependent situations, and shows how results of one set could be translated into those of the other set.

In his paper on road traffic problems, Weiss shows how the queueing problems come across in traffic are different from those of telephone communication. The reviewer feels that there is scope for applying control theory and cybernetics in highway problems. This, however, is a partially unexplored field.

Under group (iii) come the papers by Cox on statistical problems in congestion and Page on application of computers. Cox has made a detailed survey of the statistical methods on congestion problems; his paper will be particularly useful to traffic engineers and managers who would like to make meaning of observations in understanding congestion problems. Page shows how computer simulation gives an insight into various problems which do not lend to analytical solutions.

The discussions at the end of each paper are highly instructive. Each of the papers is valuable and makes an attempt to show the advances in analytical methods. There is, however, one limitation in most of the papers except those under group (iii) that attempts have been made to build a theory, in some cases on its own sake, rather than for solving some really intractable practical problems. This, however, is the limitation in the studies on queueing problems done so far. For example, we often make assumption about the Poisson stream of arrival or service epochs more as a mathematical convenience rather than a practical need. This remark does in no way undermine the qualities of the excellent studies presented in the Symposium. The get-up of the monograph is excellent.

A. Ghosal

Council of the Scientific and Industrial Research
New Delhi

Slippage Tests by Doornbos : Mathematical Centre Tracts 15, Mathematisch Centrum, Amsterdam, 1966 ; pp. 95 ; price $ 3.00.

Slippage tests are tests designed to detect 'outlying observations', or 'stragglers', or 'wildshots' as they are variously called. In essence, slippage problems are those of testing homogeneity of several samples against a special class of alternatives under which one or more of the samples deviate or 'slip' from the rest. Although such problems are of quite long standing, reasonably satisfactory solutions to the simpler forms of slippage problems have been obtained comparatively recently. In this tract the author develops one such reasonable approach and obtains useful solutions for some of the standard slippage problems.

The tract is divided into four chapters. In the first chapter the author gives a historical review of slippage problems and their solutions starting right from Benjamin Peirie (1852) and ending with Doornbos and Prins (1958). The second chapter is the most important of the four, in as much as, in it the author develops a general approach to slippage prolems with reference to the case of a single outlying sample. The general difficulty in solving such problems is that in most cases the reasonable test criterion happens to be the largest or the smallest among a set of jointly distributed random variables and hence the distribution problem becomes involved. The author solves the problem of determining the critical value to a good approximation by an application of Bonferroni's inequality subject to a condition whose validity he later proves individually in each of the special cases considered. In this way he succeeds in giving close upper and lowe rbounds to the significance level of the test. In the later part of the chapter the author considers the slippage counterparts to many of the usual tests for homogeneity. Specifically, he considers the parametric problems of homogeneity of normal means, homogeneity of scale parameters of gamma distributions, homogeneity of Poisson means and homogeneity of binomial and negative binomial probabilities. Besides he considers the slippage tests corresponding to two well-known nonparametric several samples test—the Friedman test and the Kruskal-Wallis rank-sum test. In the third chapter the author indicates an approach which might be useful in the case of more than one outlying sample, and in the concluding chapter he discusses some optimality properties of the suggested parametric slippage tests as multiple decision procedures and studies the consistency and asymptotic efficiencies of the two nonparametric slippage tests considered. At the end four tables meant to facilitate the application of some of the slippage tests are appended

This tract is a useful addition to the literature, even though it represents only one point of view. However it would have been better if the introductory review and the bibliography had been more comprehensive. In particular, one misses references to at least four important contributions on this area that have come out in recent years (notably that of Karlin and Truase in *Ann. Math. Statist.*, 1960). Another point of mild criticism is that the tract smacks a little too much like a thesis (in the attached errata one is asked to read 'tract' for 'thesis' twice!). Many mathematical details could preferably have been transferred to an appendix and the suggested procedures could have been presented in more cut and dried form with a few more numerical examples. That would have served better those interested in applications without offending those looking for the theoretical details.

S. K. Chatterjee

Department of Statistics,
Lucknow University

ABSTRACTS OF TECHNICAL REPORTS

9. **Two approximations for the distribution of double non-central beta :** PIJUSH DASGUPTA, Indian Statistical Institute, *Tech. Report* No. 8/67.

Two approximations—based on Laguerre series expansion and Jacobi series expansion have been obtained for the distribution function of the random variable $Y = \chi_2^2/(\chi_1^2+\chi_2^2)$ where χ_1^2 and χ_2^2 are independent non-central χ^2 variables with parameters $2\lambda_1$ and $2\lambda_2$, and degrees of freedom n_1 and n_2 respectively. Efficiency of these approximations have been studied numerically.

10. **An extremal problem in graph theory :** A. RAMACHANDRA RAO, Indian Statistical Institute, *Tech. Report* No. 9/67.

We consider the problem of determining the class of all connected graphs, called extremal graphs, on n vertices with m edges and having a maximum number of articulation vertices. In Theorem 5.5 we show that any extremal graph with n vertices and $m\ (\geqslant n)$ edges consists of a subgraph of one of the following types with a (possibly empty) elementary chain attached at each of its non-articulation vertices.

(1) An elementary chain (which may be a single vertex) separating a complete graph at one end and a triangle at the other end.

(2) A complete graph with another vertex joined to it by $l\ (\geqslant 2)$ edges.

(3) A graph of type (2) in which any k edges of the complete subgraph are absent, $k\ l-1$.

In a graph of the first type, the sum of the length of all free chains and the chains separating the complete graph and the triangle is $r-1$ where

$$r = \max\left\{q : m \leqslant \binom{n-q}{2}+q\right\}.$$

In a graph of the other two types, the sum of the lengths of all free chains is r.

The only extremal graph when $m = n-1$ is the elementary chain with n vertices.

The analogous problem of maximization of articulation edges is also solved.

11. **Identities involving generalised Fibonacci numbers :** MUTHULAKSHMI R. IYER, Indian Statistical Institute, *Tech. Report* No. 10/67.

Following the definition of Generalised Fibonacci Number H_n given by A. F. Horadam viz. $H_1 = p, H_2 = p+q, H_n = H_{n-1}+H_{n-2}\ n \geqslant 3$, the author has derived a number of identities involving H_n's. All the relations are given as sums upto n terms of which some of them are listed below :

Terms like $\sum_{r=1}^{n} H_{2r-1}, \sum_{r=1}^{n} H_{3r-1}, \sum_{r=1}^{n} H_{3r-1}, \sum_{r=1}^{n} H_{3r-2}$ etc.

Terms of the form $\sum_{r=1}^{n} H_{2r}^2, \sum_{r=1}^{n} H_{2r-1}^2, \sum_{r=1}^{n} H_{2r-2} H_{2r-1}$ etc.

Cubic terms like $\sum_{r=1}^{n} H_r^3, \sum_{r=1}^{n} H_{2r}^3, \sum_{r=1}^{n} H_{2r}^2 H_{2r-1}$ etc.

And lastly sums of the form

$$\sum_{r=0}^{n} rH_r, \sum_{r=0}^{n} (-1)^r rH_r, \sum_{r=0}^{n} (-1)^r H_{2r} \quad \text{etc.}$$

12. Optimal sequencing of multistage flow—shop operations: Madan Lal Mittal, Indian Statistical Institute, *Tech. Report* No. 11/67.

In this paper a method of obtaining optimal sequencing of a number of items, which have to be processed through a number of machines, is presented. It is assumed that the manufacturing time of an item on a machine is specified (i.e. non-stochastic) and the order of processing is identical for all items. The optimality criterion is the total elapsed time. The method consists in finding a lower bound on the length of all sequences in which the position of certain items is specified, proceeding with the one having the least lower bound, untill one sequence is obtained in which all the items are assigned a position and whose length is less than or equal to the lower bound on the length of all other sequences.

The Branch and Bound method considers lower bounds on the length of all sequences the first r positions of which are specified. In this paper we consider bounds on sequences whose first r_1 and /or last r_2 positions are specified. This considerably reduces the searching of the tree and thus reaches the optimum much faster than the 'Branch and Bound' algorithm.

13. Distribution of most significant digit in certain function whose arguments are random variables: A. K. Adhikari and B. P. Sarkar, Indian Statistical Institute, *Tech. Report* No. 12/67.

It is empirically well established that in large collections of numbers the proportions of entries with the most significant digit A is $\log_{10}\frac{A+1}{A}$. The porperty of the most significant digit has been studied in the present paper. It has been proved that when random numbers or their reciprocals are raised to higher and higher powers, they have log distribution of most significant digit in the limit. The property is also demonstrated in the limit by the products of random numbers as the number of terms in the product becomes higher and higher. The property is not, however, demonstrated by higher roots of the random numbers or their reciprocals in the limit. Actually there is a concentration at some particular digit. It has been shown that if X has log distribution of the most significant digit, then so does $\frac{1}{X}$ and CX, C being any constant under stronger conditions.

14. On vector variables with a linear structure and a characterization of the multivariate normal distribution: C. Radhakrishna Rao, Indian Statistical Institute, *Tech. Report* No. 13/67.

A vector random variable X is said to have a linear structure if it can be expressed as $X = \mu+AY$ where μ is a constant vector, A is a matrix and Y is a vector of independent non-degenerate random variables (called structural variables). Two structures μ_1+A_1Y and μ_2+A_2Z are said to be equivalent if one can be reduced to the other by suitable scaling and choice of location of the structural variables.

It is well known that if X is a multivariate normal variable then the structural representation is not unique both with respect to the number of structural variables and their coefficients. The converse of this proposition is proved to characterize a multivariate normal distribution. It is shown that if there exist two structural representations μ_1+A_1Y and μ_2+A_2Z of X such that no column of A_1 is a multiple of any column of A_2, then X must have a multivariate normal distribution.

Conditions under which a structural representation is unique with respect to the structural coefficients and the number of variables are investigated.

It is shown that non-uniqueness arises due to some of the structural variables having a normal distribution or having a normal component individually or in linear combinations.

Finally, a theorem is proved regarding the vector random variable X as the sum of two independent variables X_1 and X_2 where X_1 has a unique structure and X_2 is multivariate normal.

15. **On the minimal thick sets of a measure space :** S. B. RAO, Indian Statistical Institute, *Tech. Report* No. 14/67.

Let (X, S, μ) be a measure space. A set $A \subset X$ is called a thick set of (X, S, μ) if the μ-inner measure of its complement is zero. A thick set A is called a minimal thick set if no proper subset of A is a thick set. The following results are obtained.

If (X, S, μ) admits a minimal thick set A, then A is countable and μ is atomic; a finite product of minimal thick sets is a minimal thick set. Some topological lemmas are proved. If X is a complete, separable metric space without isolated points, S, the σ field generated by open subsets of X, the following conditions are equivalent.

(1) (X, S, μ) is non-atomic

(2) (X, S, μ) admits a thick set whose complement is also a thick set

(3) (X, S, μ) admits a decreasing sequence of thick sets tending to the empty set.

Some examples are given.

16. **A decomposition theorem for vector variables with a linear structure :** C. RADHAKRISHNA RAO, Indian Statistical Institute, *Tech. Report* No. 15/67.

A vector variable X is said to have a linear structure if it can be written as $X = AY$ where A is a matrix and Y is a vector of independent random variables called structural variables. In earlier papers the conditions under which a vector random variable admits different structural representations have been studied. It is shown, among other results, that complete non-uniqueness, in some sense, of the linear structure characterizes a multivariate normal variable. In the present paper we prove a general decomposition theorem which states that any variable X with a linear structure can be expressed as the sum (X_1+X_2) of two independent variables X_1, X_2 of which X_1 is non-normal and has a unique linear structure, and X_2 is a multivariate normal variable with a non-unique structure.

17. **Role of the theory of graphs in operations research :** C. RAMANUJACHARYULU, Indian Statistical Institute, *Tech. Report* No. 16/67.

The article gives an introduction to the theory graphs mainly stressing its applications to problems of operations research, Market research and linear programming. After introduction in Section 1 where the generalities on graphs are dealt with, Section 2 contains a brief historical sketch leading up to modern trends and Section 3 while introducing basic definitions also throws light on the role of graph theory in several operations research problems and related topics. New problems like constructing graph on n vertices with m edges such that no more than r edges cross at a non-vertex point and no more than such points appear on an edge are presented in course of discussions on applications.

18. **Some results on Fibonacci quaternions :** MUTHULAKSHMI R. IYER, Indian Statistical Institute, *Tech. Report* No. 17/67.

The n-th Fibonacci Quaternion Q_n is defined by the relation $Q_n = F_n+iF_{n+1}+jF_{n+2}+kF_{n+3}$ where F_n stands for the n-th Fibonacci number of the sequence 1, 1, 2, 3, 5, ... It is given by the formula $F_n = \frac{a^n-b^n}{a-b}$ where a and b are the roots of the equation $x^2-x-1=0$.

Defining analogously the n-th Lucas Quaternion T_n by $T_n = L_n+iL_{n+1}+jL_{n+2}+kL_{n+3}$ where L_n is the n-th Lucas number in the sequence 2, 1, 3, 4, 7, ..., we see that $L_n = a^n+b^n$ whereas above a, b are roots of the same quadratic equation.

In this paper some of the relations connecting F_n and L_n are presented. Then relations connecting Q_n's to F_n and L_n are derived, as also those connecting T_n's to F_n and L_n. Lastly some relations existing between the Q_n and T_n are obtained.

At present these relations seem to be of highly academic interest, but I do hope and think that there is some application of these Fibonacci and Lucas numbers.

19. **A note on an inequality for normal distribution :** C. G. KHATRI, Indian Statistical Institute, *Tech. Report* No. 18/67.

Let $x = (x_1, x_2, ..., x_p)'$ be distributed as multivariate normal with zero means and covariance matrix $V(\mathbf{x})$ · Such a law of distribution will be denoted by $x \sim N(0, V(\mathbf{x}))$. Dunn's conjecture namely

$$(1) \qquad P[|x_i| \leqslant c_i, \; i = 1, 2, ..., p] \geqslant \prod_{i=1}^{p} P[|x_i| \leqslant c_i],$$

was established recently by Khatri, Sidak and Scott by using different methods. The purpose of this note is to generalise this result in the case of convex and symmetric regions about the origins. Let $D_i(x)$ be a convex and symmetric region in x about the origin, $i = 1, 2$. Then, generalisation of (1) can be given by

$$(2) \qquad P[D_1(\mathbf{x}) \cap D_2(\mathbf{x})] \geqslant P[D_1(\mathbf{x})] \, P(D_2(x)].$$

Some applications of (2) are given with a view to illustrate the use of (2) in simultaneous confidence bounds.

20. **On range-transformation of multiple-precision numbers—I :** E. V. KRISHNAMURTHY and B. P. SARKAR, Indian Statistical Institute, *Tech. Report* No. 19/67.

An economic and efficient algorithm is proposed for range transformation of multiple precision numbers in a general radix. The rules of the algorithm are derived using error analysis.

21. **On range-transformation of multiple-precision numbers—II :** E. V. KRISHNAMURTHY and B. P. SARKAR, Indian Statistical Institute, *Tech. Report* No. 20/67.

Described in this paper is an algorithm for transforming a number such that its leading digit is earlier radix less one or the leading digit is a unity with its next digit zero. It is shown that for obtaining the required multiplier it is sufficient to have a knowledge of only the leading two digits of the given number and it is sufficient to determine the multiplier up to its leading two digits. The validity of the rules of the algorithm is proved by using diophantine analysis.

22. **Simulation of queueing problems :** A. K. ADHIKARI, Indian Statistical Institute, *Tech. Report* No. 21/67.

A method for estimating average waiting time and percentage of idle time in queueing problems by simulation has been described in the paper. The model considered is a general one where the service time parameter may depend on the arrival type. The result of a study is also presented in this paper.

23. **Polarization of nucleons elastically scattered on deuterons :** G. RAMACHANDRAN, Indian Statistical Institute, *Tech. Report* No. Phy/2/67.

The spin-orbit interaction in nucleon—nucleon scattering at low energies is studied by analysing experimental data on nucleon polarization in elastic scattering of nucleons on deuterons using the impulse approximation, which is seen to provide an elegant interpretation of the observed angular distributions of polarization from energies as low as 0.99 MeV extending upto about 25 MeV. Estimates of appropriate phase shift combinations in P, D and F partial waves are obtained using least square fits.

24. **Some properties of additive arithmetical functions :** E. M. PAUL, Indian Statistical Institute. (Not issued as a *Tech. Report*).

If an additive arithmetical function has a distribution, each value assumed by the function (and especially the value 0) belongs to the support of the distribution.

If an additive arithmetical function has a (non-uniform) distribution locally on one interval, it has a distribution; if the local distribution is proportional to Lebesgue measure, it is not known whether the result holds.

If $f_1, f_2, \ldots$ are non-negative additive functions having distributions, and α_n is a continuity point in the distribution of $f_n, n = 1, 2, \ldots$, the set of positive integers m such that simultaneously $f_1(m) \leqslant \alpha_1, f_2(m) \leqslant \alpha_2, \ldots$ has logarithmic density which is $= \lim_{n \to \infty}$ [density of $\underset{m}{E}\{f_i(m) \leqslant \alpha_i, i = 1, \ldots n\}$].