

## On Generalized Inverse of a Linear Operator

P. W. EPASINGHE\*

*Department of Statistics, Florida State University, Tallahassee, Florida 32306, U.S.A.*

(Paper accepted : 18 April 1974)

**Abstract :** The concept of a generalized inverse of a rectangular matrix is characterized by using properties of the vector spaces involved. It is then extended for linear operators from one vector space to another when the spaces involved are infinite dimensional.

The concept of a generalized inverse of a given  $m \times n$  matrix  $A$  has appeared in the literature<sup>1-4</sup> since 1920 and has been defined by Rao<sup>5-6</sup> as follows:

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

In the above definition,  $x, y$  are respectively elements of  $E^{(n)}, E^{(m)}$  which are respectively the linear spaces of complex column vectors with  $n, m$  rows.

Rao and Mitra<sup>6</sup> have derived several properties of generalized inverses, referred to as *g-inverses*, by purely algebraic techniques.

It is possible to give a geometrical characterization to a *g-inverse*. One advantage of this geometrical characterization is that the results can immediately be extended to linear operators from one space to another space, when the spaces involved are not necessarily finite dimensional. We handle first, the case considered by Rao and Mitra, and then proceed to show how the concept can be extended to linear operators.

The given  $m \times n$  matrix  $A$  may be considered as a linear mapping of  $E^{(n)}$  into  $E^{(m)}$ . Let  $N(A)$  be the null space of  $A$  i.e.  $N(A)$  is the set of all  $x \in E^{(n)}$  such that  $Ax = 0$ , the additive identity of  $E^{(m)}$ . Then we have the following elementary result :—

**Lemma 1.** The rank of  $A$  is  $n$  if the dimension of  $N(A)$  is zero.

In this case, when  $N(A)$  is zero dimensional,  $Ax = y$  will have a unique solution for  $x$ , for any  $y$  for which this is consistent. This unique solution is given by  $x = A_L^{-1}y$  where  $A_L^{-1} = (A^*A)^{-1}A^*$  is a left inverse<sup>7</sup> of  $A$ .

\*On sabbatical leave from the University of Sri Lanka, Vidyodaya Campus, Nugegoda, on a Fulbright Post Doctoral Fellowship.

Let us consider the case when  $N(A)$  is not zero dimensional. If its dimension is  $r$ , then  $0 < r \leq n$ , since  $N(A)$  is a subspace of  $E^{(n)}$ . Let  $\{x_1, x_2, \dots, x_r\}$  be a basis for  $N(A)$ . Then  $E^{(n)}$  has a basis of the form  $\{x_1, x_2, \dots, x_r, x_1^{(1)}, x_2^{(1)}, \dots, x_{n-r}^{(1)}\}$ , with obvious modifications if  $r = n$ . In the following discussion we will always assume that  $r < n$ , because if  $r = n$ , the only  $y$  for which  $Ax = y$  is consistent is  $y = 0$ , and any  $x \in E^{(n)}$  is a solution. The following results are then obtained:—

*Lemma 2.* The set  $\{Ax_1^{(1)}, Ax_2^{(1)}, \dots, Ax_{n-r}^{(1)}\}$  is linearly independent.

$$\begin{aligned} \sum_{i=1}^{n-r} \lambda_i Ax_i^{(1)} = 0 &\Rightarrow A \sum_{i=1}^{n-r} \lambda_i x_i^{(1)} = 0 \\ \Rightarrow \sum_{i=1}^{n-r} \lambda_i x_i^{(1)} \in N(A) &\Rightarrow \sum_{i=1}^{n-r} \lambda_i x_i^{(1)} = \sum_{i=1}^r \mu_i x_i \\ \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_{n-r} = \mu_1 = \mu_2 = \dots = \mu_r = 0 \end{aligned}$$

since  $\{x_1, x_2, \dots, x_r, x_1^{(1)}, x_2^{(1)}, \dots, x_{n-r}^{(1)}\}$  is a linearly independent set.

*Lemma 3.* The set  $\{Ax_1^{(1)}, Ax_2^{(1)}, \dots, Ax_{n-r}^{(1)}\}$  is a basis for the range  $M(A)$  of  $A$ .

If  $y \in M(A)$ , then  $\exists x \in E^{(n)}$  such that  $Ax = y$ .

$$\text{But } x = \sum_{i=1}^r \lambda_i x_i + \sum_{i=1}^{n-r} \mu_i x_i^{(1)} \text{ for suitable scalars } \lambda_1, \lambda_2, \dots, \lambda_r, \mu_1, \mu_2, \dots, \mu_{n-r}.$$

$$\therefore y = \sum_{i=1}^{n-r} \mu_i (Ax_i^{(1)}) \text{ since } Ax_1 = Ax_2 = \dots = Ax_r = 0.$$

*Lemma 4.* If  $R(A)$  is the subspace (of  $E^{(n)}$ ) spanned by the set  $\{x_1^{(1)}, x_2^{(1)}, \dots, x_{n-r}^{(1)}\}$ , then the linear mapping  $A : R(A) \rightarrow M(A)$  is one-one and onto.

Hence there is a unique linear mapping  $A^- : M(A) \rightarrow R(A)$  which is the inverse of the mapping  $A : R(A) \rightarrow M(A)$ . Now, we have the following theorem:—

*Theorem 1.* If  $y$  is any vector for which  $Ax = y$  is consistent, then  $x = A^-y$  is a solution of the equation  $Ax = y$ , and therefore  $A^-$  is a g-inverse of  $A$ .

At this stage it is worthwhile to note that, although the linear mapping  $A^-: M(A) \rightarrow R(A)$  defined is unique, its matrix representation is not. For instance, if  $[A^-]_{n \times m}$  is any  $n \times m$  matrix satisfying the requirement, and  $[N]_{n \times m}$  is any  $n \times m$  matrix which maps  $M(A)$  into the null vector of  $E^{(a)}$ , then  $[A^-]_{n \times m} + [N]_{n \times m}$  is also a matrix representation of the mapping  $A^-$ .

It is easily seen that each extension of a basis for  $N(A)$  to form a basis for  $E^{(a)}$  provides us with one g-inverse. Conversely, suppose  $B$  is a  $n \times m$  matrix such that, for each  $y \in M(A)$ ,  $Ax = y$  is satisfied by  $x = By$ . Let  $\{y_1, y_2, \dots, y_{n-r}\}$  be a basis for  $M(A)$ . Then  $\{z_i = By_i, i = 1, 2, \dots, n-r\}$  is a set of  $n-r$  linearly independent vectors belonging to  $E^{(a)}$ . Further  $\{x_1, x_2, \dots, x_r, z_1, z_2, \dots, z_{n-r}\}$  is set of  $n$  linearly independent vectors belonging to  $E^{(a)}$  and hence forms a basis for  $E^{(a)}$ . Therefore,  $\{x_i \mid i = 1, 2, \dots, r\} \cup \{z_i \mid i = 1, 2, \dots, n-r\}$  is an extension of the basis  $\{x_i \mid i = 1, 2, \dots, n\}$  of  $N(A)$  to form a basis for  $E^{(a)}$ . This, as we know, gives rise to a unique g-inverse of  $A$ , and  $B$  is one of its matrix representations.

The possible non-uniqueness of the matrix representation of  $A^-$  may appear to give rise to a certain amount of confusion. However, the following result<sup>6</sup> gives a simple relationship among all the matrix representations.

*Lemma 5.* If  $[A^-]_{n \times m}$  is one matrix representation of a g-inverse of  $A$ , then every matrix representation of every g-inverse of  $A$  is of the form  $[A^-]_{n \times m} + [N]_{n \times m}$  where  $N$  is an  $n \times m$  matrix satisfying the condition  $ANA = 0$  as a  $n \times n$  matrix equation. Conversely, for every such  $N$ ,  $[A^-]_{n \times m} + [N]_{n \times m}$  is a generalized inverse of  $A$ .

This follows from Lemma 2.2.1 of Rao and Mitra.<sup>6</sup>

What all this shows is that we have generated the set of all g-inverses of the  $m \times n$  matrix  $A$ . Each such g-inverse has an  $n \times m$  matrix representation and hence may be considered as a linear mapping of  $E^{(m)}$  into  $E^{(n)}$ . Since  $M(A)$  is an  $n-r$  dimensional subspace of  $E^{(m)}$ , and each  $A^-$  maps  $M(A)$  onto an  $n-r$  dimensional subspace of  $E^{(n)}$ , it is clear that the rank of  $A^-$  is not less than  $n-r$  which is the rank of  $A$ . Hence  $\text{rank } (A^-) \geq \text{rank } A$ .

Now we attempt to extend the above characterization to a general linear operator  $A$  which maps a given linear vector space  $V$  into another linear vector space  $W$ , both spaces assumed to be infinite dimensional to avoid unnecessary complications in notation. The following definitions and theorems<sup>8</sup> pertaining to infinite dimensional linear vector spaces are quoted for completeness.

*Definition 2.* A set  $\{x\}$  of vectors of an infinite dimensional linear vector space is said to be *linearly independent* if every finite subset of  $\{x\}$  is linearly independent. Otherwise, the set  $\{x\}$  is said to be *linearly dependent*.

*Definition 3.* A linearly independent set  $\{x\}$  of vectors of a linear vector space  $V$  is called a *Hamel basis*, if given any  $v \in V$ ,  $\exists$  a finite subset  $\{x_1, x_2, \dots, x_k\}$  of  $\{x\}$  and scalars  $\lambda_1, \lambda_2, \dots, \lambda_k$ , such that

$$v = \sum_{i=1}^k \lambda_i x_i$$

*Lemma 6.* Every linear space has a Hamel basis. Each vector is a unique linear combination of a finite number of vectors of the Hamel basis.

*Lemma 7.* If  $U$  is a subspace of  $V$ , and  $\{x\}$  is a Hamel basis for  $U$ , then  $V$  has a Hamel basis  $\{y\}$  such that  $\{x\} \subset \{y\}$ . In other words, every linearly independent set of vectors of  $V$  can be extended to form a Hamel basis for  $V$ .

Now, let us consider a linear operator  $A$  which maps  $V$  into  $W$ ,  $V$  and  $W$  both being infinite dimensional. The set of all  $v \in V$  such that  $Av = \theta$ , the additive identity of  $W$ , is a linear subspace  $N(A)$  of  $V$ . We will call  $N(A)$ , the *null space* of  $A$ .

*Lemma 8.* If the null space of  $A$  is zero dimensional, then the equation  $Ax = y$  where  $x \in V, y \in W$  has a unique solution for each  $y$  for which  $Ax = y$  is consistent.

If  $M(A)$  is the range of  $A$ , then the mapping  $\Lambda : V \rightarrow M(A)$  is one-one and onto. Hence, there is a unique inverse mapping  $\Lambda^{-1} : M(A) \rightarrow V$ . Therefore, given  $y \in M(A)$ ,  $x = \Lambda^{-1}y$  is the unique solution of  $Ax = y$ .

Let us now consider the case when  $N(A)$  is not zero dimensional. Here again, if  $N(A) = V$ , then the only  $y$  for which  $Ax = y$  is consistent is  $y = \theta$ , and therefore for an arbitrary linear operator  $B$  from  $W$  into  $V$ ,  $x = By$  is a solution of the equation of the equation  $Ax = y$ . Hence, in the following discussion, we assume that  $N(A)$  is a proper subspace of  $V$ .

Let  $\{x\}$  be a Hamel basis for  $N(A)$ , which exists by Lemma 5. Then by Lemma 6,  $V$  has a Hamel basis of the form  $\{x\} \cup \{x^{(i)}\}$ .

*Lemma 9.*  $\{Ax^{(i)}\}$  is a linearly independent set.

If  $\{Ax_1^{(i)}, Ax_2^{(i)}, \dots, Ax_n^{(i)}\}$  is any finite subset of elements of the above set, then,

$$\begin{aligned} \sum_{i=1}^n \lambda_i Ax_i^{(i)} = \theta &\implies \Lambda \sum_{i=1}^n \lambda_i x_i^{(i)} = \theta \\ \implies \sum_{i=1}^n \lambda_i x_i^{(i)} \in N(A) &\implies \exists \text{ a finite subset} \end{aligned}$$

$\{x_1, x_2, \dots, x_m\}$  of elements of  $\{x\}$  and scalars  $\mu_1, \mu_2, \dots, \mu_m$  such that

$$\sum_{i=1}^n \lambda_i x_i^{(1)} = \sum_{i=1}^m \mu_i x_i \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = \mu_1 = \mu_2 = \dots = \mu_m = 0,$$

by the linear independence of the set  $\{x\} \cup \{x^{(1)}\}$ .

*Lemma 10.*  $\{Ax^{(1)}\}$  is a Hamel basis for  $M(A)$ , the range of  $A$ .

Proof as in Lemma 3, with obvious modifications in notation.

*Lemma 11.* If  $R(A)$  is the subspace (of  $V$ ) spanned by the linearly independent set  $\{x^{(1)}\}$ , then the linear mapping  $A : R(A) \rightarrow M(A)$  is one-one and onto.

Hence, there exists a unique linear operator  $A^- : M(A) \rightarrow R(A)$  which is the inverse mapping of  $A : R(A) \rightarrow M(A)$ .

*Definition 4.*  $A^-$  defined above is called a g-inverse of  $A$ .

It is seen that each extension of a basis of  $R(A)$  to form a basis of  $V$ , provides us with one g-inverse of  $A$ . The last result we have is as follows :

*Theorem 2.* If  $A : V \rightarrow W$  is a linear mapping of a linear vector space  $V$  into a linear vector space  $W$ , then  $\exists$  a linear operator  $A^-$  such that, for each  $y \in W$  for which  $Ax = y$  is consistent,  $x = A^-y$  is a solution of the equation  $Ax = y$ .

**Acknowledgements**

The author wishes to acknowledge with thanks the award of a grant by the State Department of the United States for advanced study and research sponsored under the Fulbright-Hays Act. He is also grateful to Professor Ralph A. Bradley for hospitality at the Statistics Department of the Florida State University, where this work was done, while being supported by funds provided under the above grant.

**References**

1. MOORE, E. H. (1920) *Bull. Am. Math. Soc.* 26 : 394-395.
2. MOORE, E. H. (1935) *General analysis*. Philadelphia : American Philosophical Society.
3. PENROSE, E. (1955) *Proc. Camb. Phil. Soc. Math. & Phys. Sci.* 51 : 406-413.
4. PENROSE, R. (1956) *Proc. Camb. Phil. Soc. Math. & Phys. Sci.* 52 : 17-19.
5. RAO, C. R. (1962) *J. R. Statist. Soc. Ser. B.* 24 : 152-158.
6. RAO, C. R. & MITRA, S. K. (1971) *Generalized inverse of matrices and its applications*, p. 20. New York : John Wiley.
7. RAO, C. R. & MITRA, S. K. (1971) *Ibid.*, p. 19.
8. WILANSKY, A. (1964) *Functional analysis*. New York : Blaisdell Publishing Co.