Application Of Generalized Inverse Of A Matirx To Models Not Of Full Rank

Adetunde, I . A.
University of Mines & Technology
Department of Mathematics
Tarkwa Ghana

Oladejo, Nathaniel
University for Developments Studies
Department of Mathematics and Computer Science
Navrongo. Ghana

Apio, Dominc
University for Developments Studies
Department of Mathematics and Computer Science
Navrongo. Ghana

adetunde@googlemail.com

ABSTRACT: This paper considered the application of generalized inverse of a matrix to models not of full rank. In the recent paper (1), On the generalized inverse of a matrix, the generalized inverse matrix was applied to solution of systems of equations that are linearly dependent and unbalanced. This paper is an extensive study of (1). It deals with the application of generalized inverse of a matrix to models that are not full rank. [Researcher. 2009;1(2):41-53]. (ISSN: 1553-9865).

KEYWORDS: Generalized Inverse of a matrix, linear models, least square, Estimation, Variance, Full rank partitioning

INTRODUCTION

The work presented in this paper is an extension of the earlier paper [1]. Here the generalized inverse of a matrix is applied to models which are not of full rank in nature.

Generalized inverse of a matrix is a research area in mathematical statistics. We are all aware of the fact that every non-singular matrix A has a unique inverse denoted by A such that $AA^{-1} = A^{-1}A = I$ where I is the identity matrix. That is a matrix has an inverse only if it is square and if only if it is non-singular or in other words, if its columns (rows) are linearly independent.

In recent years needs have been felt in numerous areas of applied mathematics for some kind of partial inverse of a matrix that is singular or even rectangular, hence the beginning of the use of generalized inverse. The most familiar applications of matrices is to the solution of systems of simultaneous linear equation, and the application of generalized inverse is no exception. Generalized inverse is also applied to least squares estimate (LSE) in the study of linear models.

Various methods exist for solving systems of simultaneous linear equation; some of them are: (elimination method, row reduction method, backward substitution method etc). Require that the set of linear equations be linearly independent. What if the system of equations is linearly dependent? Generalized inverse is able to solve linearly dependent and unbalance system of equations.{See Paper[1]}

Generalized inverse are of great importance in its general application to non-square and square singular matrices. In the case that A is non-singular $G = A^{-1}$ and G is unique.
The fact that A has a generalized inverse even if it is singular or rectangular has particular applications in the problem of solving equations like

$$AX = Y$$

More- over, generalized inverse are of great importance in the study of linear models where least square estimate often leads to equation of the form

$$X^1 Xb = X^1 Y$$

This has to be expressed in the form

$$b = \left(X^1 X\right)^{-1} X^1 Y$$

But if $X^1 X$ is singular then $\left(X^1 X\right)^{-1}$ does not exist hence the use of generalized inverse to solve such system of equations is needed, which is the main objective of this study. In mathematics, a generalized inverse or pseudoinverse of a matrix *A* is a matrix that has some properties of the inverse matrix of *A* but not necessarily all of them. The term "the pseudoinverse" commonly means the Moore-Penrose pseudoinverse.

The purpose of constructing a generalized inverse is to obtain a matrix that can serve as the inverse in some sense for a wider class of matrices than invertible ones. Typically, the generalized inverse exists for an arbitrary matrix, and when a matrix has inverse, then its inverse and the generalized inverse are the same. Some generalized inverses can be defined in any mathematical structure that involves associative multiplication, that is, in a semi group.

The various kinds of generalized inverses include

- one-sided inverse, that is left inverse and right inverse
- Drazin inverse
- Group inverse
- Bott–Duffin inverse (in German)
- Moore-Penrose pseudoinverse

1.2 HISTORICAL BACKGROUND OF GENRALIZED INVERSE MATRIX

The concept of a generalized inverse seems to have been first mention in print in 1903 by Fredholm, where a particular generalized inverse called by him pseudo inverse as an integral operator was given. Several investigations have concerned themselves with the Generalized inverse matrices, notably among them were: Hurwitz (1912), He characterized all pseudo inverse and used the finite dimensionality of null operators of Fredholm operators, already implicit in Hilbert's discussions in 1904 of generalized Green functions were consequently studied by numerous authors, in particular, Myller (1906), Westfall (1909), Bounitzky (1909), Elliott (1928) , Reid (1931). Bjerhanmer (1951), Penrose (1951)

Relevant publications are the work done by Moore (1920), Siegel (1937), Tseng, Murray and Von Neumann (1936), Alkinson (1950), Adetunde et al; (2008).

2. ALGORITHM FOR THE GENERALIZED INVERSE OF A MATRIX

An algorithm for finding the generalized inverse of a matrix is as follows, according to Adetunde et al; (2008)

Step 1: in A of rank r, find any non-singular minor of order r call it M

Step 2: invert M and transpose the inverse (M)

Step3: in A replace each element of M by the corresponding element of (M)

That is a = M the (s,t) element of m, then replace a b M, the (t,s) element of M equivalent to the (s,t) element of the transpose of M

Step4: replace all the other elements of A by zero

Step 5: transpose the resulting matrix and the result is G a generalized inverse of A

2.1 PROPERTIES OF GENERALIZED INVERSE OF A MATRIX

If G is a generalized inverse of A then

- AGA = A
- G is not unique
- G is of order m x n if A is of oder n x m

If G is a generalized inverse of XX then

- G is also a generalized inverse of XX
- XGXX = X; that is , GX is a generalized inverse of X
- XGX is invariant to G
- XGX is symmetric, whether G is or not

3 APPLICATION OF GENERALIZED INVERSE TOMODELS NOT OF FULL RANK

The model we shall be dealing with is

$$Y = Xb + e$$

Where Y is an N x 1 vector of observations y_i

B is a P x 1 vector of parameters X is an N x P matrix of known values (in most cases 0's and 1's) and e is a vector of random error terms.

The following assumptions are made

$$e \approx (0, \sigma^2 I)\; and\;\; Y \approx \left(Xb, \sigma^2\, I\right)$$

$\Rightarrow$ The random errors are distributed normally with a zero mean and constant variance $\sigma^2\; I\; and\;\; Y$ is also distributed normally with a mean of Xb and a constant variance $\sigma^2\; I.$ The normal equation corresponding to the model is given as

$$Y = Xb + e$$

which can be derived by the least squares method, to get

$$X^1\ Xb = X^1 Y$$

Example

In an experiment to estimate the effect of type of plant on the weight of the maize fruit four different maize plants given the same condition recorded the following weight of its fruit at harvest as

Weight of 10 plants	Type 1	Type 2	Type 3	Type 4
	62	80	62	60
	71	75	75	
	83	45		
	90			
Total	306	200	137	60

To estimate the effect of the type of plant on the weight of plant we assume that the observation $y_{i,j}$ is the sum of four parts

$$Y_{i,j} = \nu + \alpha_I + e_{i,j}$$

Where

ν is the population mean of the weight of plant

α_I is the effect of the type I on weight

$e_{i,j}$ is the random error term peculiar to the observation $y_{i,j}$

To develop the normal equations, we write down 10 observations in terms of the equation of the model

$$62 = y_{11} = \nu + \alpha_I + e_{11}$$
$$71 = y_{12} = \nu + \alpha_I + e_{12}$$
$$83 = y_{13} = \nu + \alpha_I + e_{13}$$
$$90 = y_{14} = \nu + \alpha_I + e_{14}$$
$$80 = y_{21} = \nu + \alpha_I + e_{21}$$
$$75 = y_{22} = \nu + \alpha_I + e_{22}$$
$$45 = y_{23} = \nu + \alpha_I + e_{23}$$
$$62 = y_{31} = \nu + \alpha_I + e_{31}$$
$$75 = y_{32} = \nu + \alpha_I + e_{32}$$
$$60 = y_{41} = \nu + \alpha_I + e_{41}$$

This is written in matrix form as

$$\begin{pmatrix} 62 \\ 71 \\ 83 \\ 90 \\ 80 \\ 75 \\ 45 \\ 62 \\ 75 \\ 60 \end{pmatrix} = \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{41} \end{pmatrix} = \begin{pmatrix} 1\,1\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,0\,1\,0\,0 \\ 1\,0\,1\,0\,0 \\ 1\,0\,1\,0\,0 \\ 1\,0\,0\,1\,0 \\ 1\,0\,0\,1\,0 \\ 1\,0\,0\,0\,1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{14} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{31} \\ e_{32} \\ e_{41} \end{pmatrix}$$

$$Y \quad = \quad X \qquad b \quad + \quad e$$

Y is the vector of observations, X is the incidence matrix and b is the vector of parameters to be considers

The normal equations corresponding to the model

$Y = Xb + e$ can be derived by least square to give

$$X^1 Xb = X^1 Y$$

$$X^1 X = \begin{pmatrix} 1111111111 \\ 1111000000 \\ 0000111000 \\ 0000000110 \\ 0000000001 \end{pmatrix} \begin{pmatrix} 1\,1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0\,0 \\ 1\,0\,1\,0\,0\,0 \\ 1\,0\,1\,0\,0\,0 \\ 1\,0\,1\,0\,0\,0 \\ 1\,0\,0\,1\,0\,0 \\ 1\,0\,0\,1\,0\,0 \\ 1\,0\,0\,0\,1\,0 \end{pmatrix} = \begin{pmatrix} 10 & 4 & 3 & 2 & 1 \\ 4 & 4 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$X^1Y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 62 \\ 71 \\ 83 \\ 90 \\ 80 \\ 75 \\ 45 \\ 62 \\ 75 \\ 60 \end{pmatrix} = \begin{pmatrix} 703 \\ 306 \\ 200 \\ 137 \\ 60 \end{pmatrix}$$

Matrix X^1X has determinant equal to zero and hence not of full rank, therefore matrix X^1X has no unique inverse, hence the equation cannot be express as

$$b = \left(X^1X\right)^{-1}\left(X^1Y\right)$$

since $\left(X^1X\right)^{-1}$ does not exist.
This implies that the normal equation has no unique solution. To get one of the solution, we need to find any generalized inverse G of $\left(X^1X\right)$ and write the corresponding solution as

$$b^0 = GX^1Y$$

where G is a generalized inverse of X^1X

the notation b^0 *and not* b used in equation emphasizes that what is derived by solving is only a solution to the equations and not an estimator of b

choosing

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$b^0 = GX^1Y$$

$$\begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 703 \\ 306 \\ 200 \\ 137 \\ 60 \end{pmatrix} = \begin{pmatrix} 0 \\ 76.5 \\ 66.7 \\ 68.5 \\ 60 \end{pmatrix}$$

The expectation of b^0 is given as

$E(b^0) = GX^1 E(Y)$

$E(Y) = Xb$

$\therefore\ E(b^0) = GX^1 Xb$

$E(b^0) = Hb$

Where $H = GX^1 X$ hence b^0 is an unbiased estimator of Hb but not of b

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 10 & 4 & 3 & 2 & 1 \\ 4 & 4 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The variance of b^0 is given as

$$Var(b^0) = Var\left(GX^1 Y\right)$$

$$= GX^1 Var(Y) XG^1$$

$$= GX^1 XG^1 \sigma^2$$

For a full rank model $Var(b) = (X^1X)^{-1}\sigma^2$, by an appropriate choice of G, $GX^1 XG^1\sigma^2$ can reduce further to $G\sigma^2$

Estimating E(y)

Corresponding to the vector of observations y, we have the vectors of estimated expected values $E(\hat{y})$.

$$E(\hat{y}) \equiv \hat{y} = Xb^0 = XGX^1Y$$

This vectors is invariant to the choice of whatever generalized inverse of X^1 Xis used for G, because XGX^1

is invariant. This means that no matter what solution of the normal equations is used for b^0 the vector $\hat{y} = XGX^1Y$ will always be the same.

$$\hat{Y} = Xb$$

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{41} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 76.5 \\ 66.7 \\ 68.5 \\ 60.0 \end{pmatrix} = \begin{pmatrix} 75.6 \\ 75.6 \\ 75.6 \\ 75.6 \\ 66.7 \\ 66.7 \\ 66.7 \\ 68.5 \\ 68.5 \\ 60.0 \end{pmatrix}$$

$\hat{y}$ *is the vector of* exp*ected values.*.

To demonstrate the invariance of $\hat{y}$ to the choice of G. Consider

$$G = \begin{pmatrix} 1 & -1 & -1 & -1 & 0 \\ -1 & 5/4 & 1 & 1 & 0 \\ -1 & 1 & 4/3 & 1 & 0 \\ -1 & 1 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$H = GX^1X$, hence we have

$$H = \begin{pmatrix} 1 & -1 & -1 & -1 & 0 \\ -1 & 5/4 & 1 & 1 & 0 \\ -1 & 1 & 4/3 & 1 & 0 \\ -1 & 1 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 10 & 4 & 3 & 2 & 1 \\ 4 & 4 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$b^0 = GX^1Y = \begin{pmatrix} 1 & -1 & -1 & -1 & 0 \\ -1 & 5/4 & 1 & 1 & 0 \\ -1 & 1 & 4/3 & 1 & 0 \\ -1 & 1 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 703 \\ 306 \\ 200 \\ 137 \\ 60 \end{pmatrix} = \begin{pmatrix} 60 \\ 16.5 \\ 6.7 \\ 8.5 \\ 0 \end{pmatrix}$$

$$\hat{Y} = Xb^0$$

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{41} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 60 \\ 16.5 \\ 6.7 \\ 8.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 75.6 \\ 75.6 \\ 75.6 \\ 75.6 \\ 66.7 \\ 66.7 \\ 66.7 \\ 68.5 \\ 68.5 \\ 60 \end{pmatrix}$$

From the above results, it demonstrates that $\hat{y}$ is always the same no matter the G used.
The sum of squares regression is also invariant to the choice of G

4 PARTITIONING THE TOTALL SUM OF SQUARES

Partitioning the total sum of square for the full rank model is the same for the model not of full rank. The only difference is that there is utility in corrected sums of squares and products of the x – variables.

$SS_T = Y^1Y - N\bar{Y}^2$

$SS_R = (b^0)^1 X^1Y - N\bar{Y}^2$

$SS_E = SS_T - SS_R$

$$Y^1Y = (62\ 71\ 83\ 90\ 80\ 75\ 45\ 62\ 75\ 60)\begin{pmatrix}62\\71\\83\\90\\80\\75\\45\\62\\75\\60\end{pmatrix} = 50993$$

$Y^1\ Y = 50993$

$N = 10,\ \bar{Y} = 70.3$

$SS_T = 50993 - 10(70.3)^2$

$= 50993 - 10(4942.09)$

$= 50993 - 49420.9$

$SS_T = 1572.1$

$$b^0X^1\ Y = (\ 60\ \ 16.5\ \ 6.5\ \ 8.5\ \ 0)\begin{pmatrix}703\\306\\200\\137\\60\end{pmatrix} = 49733.5$$

$SS_R = 49733.5 - 49420.9$

$SS_R = 312.6$

SS_R is invariant to the choice of G, to show the invariance of SS_R , we consider

$$b^0 = \begin{pmatrix} 0 \\ 76.5 \\ 66.7 \\ 68.5 \\ 60 \end{pmatrix}$$

$$b^0 X^1 Y = (0 \quad 76.5 \quad 66.7 \quad 68.5 \quad 60) \begin{pmatrix} 703 \\ 306 \\ 200 \\ 137 \\ 60 \end{pmatrix} = 49733.5$$

$SS_T = 49733.5 - 49420.9$

$= 312.6$

Since SS_R is the same, no matter the G used we say that SS_R is invariant to G

$SS_E = SS_T - SS_R$

$= 1572.1 - 312.6$

$= 1259.5$

Test the hypothesis

$H_0 : Xb = 0$

$H_1 : Xb \neq 0$

For the full rank model the test hypothesis is

$H_0 : b = 0$

$H_1 : b \neq 0$

But for a model not of full rank b is not estimable, hence the hypothesis

$H_0 : b = 0$

$H_1 : b \neq 0$

Cannot be tested because b is a non-estimable function

ANOVA table

Source	Df	SS	MS	F
Regression Residual Total	3 6 9	312.6 1259.5 1572.1	104.2 209.9	0.496

R^2 = 312.6/1572.1 = 0.199

The total variation explained by the model is 19.9%, the overall model is not significant, meaning that the weight of the maize fruit does not depend on the type of maize plant.

SAS OUTPUT

CONCLUSION

In this paper, the method of generalized inverse had been applied on linear models which is not of full rank. Evidence has shown from our result that generalized inverse can not be overlooked since it plays a very important role in models not of full rank. Most importantly, the use of generalized inverse of a matrix enables us to solve systems of linear equations that are unbalance and linearly dependent easily.

REFERENCES

(1) Bing Zheng and R. B. Bapat, (2004): *Generalized inverse A(2)T,S and a rank equation*,Applied Mathematics and Computation 155 (2004) 407-415

(2) S. L. Campbell and C. D. Meyer, (1991): *Generalized Inverses of LinearTransformations*, Dover

(3) Adi Ben-Israel and Thomas N.E. Greville, (2003): *Generalized inverses. Theory and applications*. 2nd ed. New York, NY: Springer, 2003.

(4) C. Radhakrishna Rao and Sujit Kumar Mitra,(1971): *Generalized Inverse Of Matrices and its Applications*, John Wiley & Sons New York, 1971, 240 p.

(5) Ben-Israel, A. and Greville, T. N. E.(1977): Generalized Inverses: Theor and Applications. New York: Wiley, 1977.

(6) Campbell, S. L. and Meyer, C. D. Jr.(1991): Generalized Inverse of Linear Transformations. New York: Dover

(7) Lawson, C. and Hanson, R.(1974): Solving Least Squares Problems. Englewood Cliffs, NJ: Prentice-Hall

(8) Penrose, R.(1955): "A Generalized Inverse for Matrices." *Proc. Cambridge Phil. Soc.* **51**, 406-413,

(9) Rao, C. R. and Mitra, S. K.(1971): Generalized Inverse of Matrices and Its Applications. New York: Wiley

(10) Adetunde, I.A and Dominic A (2009):On the generalized inverse of a matrix. Journal of Advances and Applications in Discrete Mathematics. http:www.pphmj.com

2/18/2009