A NOTE ON THE USE OF GENERALIZED INVERSE OF MATRICES IN STATISTICS

H. ÖZDEN

1. INTRODUCTION

It is well known that if A is a non-singular matrix, then there exists a matrix G, such that $AG = GA = I$ which is called the inverse of A and denoted by A^{-1}. Moore extended the notion of inverse to singular matrices in 1920 [3], and published some of his results at some length in 1935 [4]. Another definition of inverse G of A, came from Penrose in 1955 [5] satisfying more conditions than the definition of Moore. More and more statisticians are becoming interested in this new concept.

Carter and Myers (1972) [1], Speed (1974) [7], Mazunder (1980) [2], Searle (1984) [6] are some of them.

2. DEFINITIONS AND PRELIMINARIES

Definition 1. Let A be an $m \times n$ matrix with arbitrary rank. A generalized inverse of A is an $n \times m$ matrix G such that $X = GY$ is a solution of consistent equations $AX = Y$ [8].

Additional conditions are usually imposed on a generalized matrix G so as to obtain useful results. For instance, G which was defined by Penrose is more complicated than G defined by Moore.

Since we do not need more complicated G here, we will be content with the following inverse which we call a g-inverse.

Definition 2. A g-inverse of A of order $m \times n$ is a matrix G of order $n \times m$ such that $AGA = A$.

3. PROPERTIES OF G-INVERSE

Let G be a g-inverse of A, then with $H = GA$,

i) $r(H) = tr(H) = r(A)$,

ii) $r(G) > r(A)$,

iii) $H^2 = H$,

iv) If G is a g-inverse of A then G' will be a g-inverse of A'.

4. SOLVING LINEAR EQUATIONS USING GENERALIZED INVERSES

Theorem 1. Let A be of order $m \times n$ and G be any g-inverse of A. Further let $H = GA$. Then the following hold [8]:

i) A general solution of the homogeneous equation $AX=0$ is $X=(H-I)Z$ where Z is an arbitrary vector,

ii) A general solution of consistent non homogeneous equation $AX = Y$ is

$$X = GY + (H - I)Z$$

where Z is an arbitary vector,

iii) $Q'X$ has a unique value for all solutions of $AX = Y$ iff

$$H'Q = Q.$$

5. MODEL FOR ANALYSIS OF VARIANCE

The mathematical model for one-way variance analysis is

$$Y = X\beta + \varepsilon,$$

where Y is an $m \times 1$ vector of observation; β is an $m \times n$ dummy matrix, mostly called as a design matrix; β is an $n \times 1$ vector of parameters; ε is an $m \times 1$ vector of error terms such that $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = o^2 I$.

The normal equations of the model is

$$(X'X)\widehat{\beta} = X'Y$$

where $\widehat{\beta}$ is the estimator of β. Hence the model X is singular and so $X'X$ is.

Theorem 2. Let G be any g-inverse of $(X'X)$. Then the following hold:

i) **G is also a g-inverse of (X′ X),**

ii) $XGX'X = X$, in other words GX' is a g-inverse of X,

iii) XGX' is invariant for G,

iv) XGX' is both symmetric and idempotert.

6. ESTIMATORS AND THEIR VARIANCES

Using Theorem 1, the estimator of β is written as

$$\widehat{\beta} = GX'Y + (H - I)\,Z, \qquad \text{(i)}$$

then it is clear that the expected value of $\widehat{\beta}$ is

$$\begin{aligned} E(\widehat{\beta}) &= E\,[GX'Y + (H - I)\,Z] \\ &= GX'\,E(Y) + (H - I)\,Z, \\ &= GX'\,X\beta + (H - I)\,Z \text{ since } E(Y) = X\beta, \qquad (2) \\ &= H\beta + (H - I)\,Z \text{ since } H = GX'X. \end{aligned}$$

$$\rightarrow E(\widehat{\beta}) \neq \beta \;\rightarrow\; \widehat{\beta} \text{ is a bias estimator of } \beta.$$

Coming to the variance of $\widehat{\beta}$, for Z is an arbitrary constant it does not affect the variance of $\widehat{\beta}$. Then by the definition of variance

$$\begin{aligned} V(\widehat{\beta}) &= E\,[\widehat{\beta} - E(\widehat{\beta})]\,[\widehat{\beta}' - E(\widehat{\beta}')]. \text{ Using (1) and (2)} \\ &= E\,[GX'Y - H\beta]\,[Y'XG' - \beta'H']. \text{ Since } H = GX'X \\ &= E\,[GX'Y - GX'X\beta]\,[Y'XG' - \beta'X'XG'] \\ &= GX'E\,(Y - X\beta)\,(Y' - \beta'X')\,XG' \\ &= GX'E\,(\varepsilon\varepsilon')\,XG' \\ &= GX'XG'\,o^2 \\ &= G\,\sigma^2. \qquad (3) \end{aligned}$$

Which is a similar result to the regression situation when $X'X$ is non-singular.

Estimating the error variance

$$\begin{aligned} SSE &= (Y' - \widehat{Y}')\,(Y - \widehat{Y}) \\ &= (Y' - Y'XGXY)\,(Y - XGX'Y) \\ &= Y'\,(I - XGX')\,Y. \qquad (4) \end{aligned}$$

Now G was not unique but XGX' was invariant. Hence no matter which G is used (4) will take the same value. Furthermore $X(H - I)$ is null,

$$\begin{aligned} SSE &= Y'Y - Y'X\widehat{\beta} \\ &= Y'Y - \widehat{\beta}'X'Y. \qquad (5) \end{aligned}$$

Although G is not unique XGX' is, hence (5) will be invariant and unique.

In order to obtain the expected value of SSE, it will be enough to put $Y = X\beta + \varepsilon$ in (5) and remember that $(I - XGX')$ is an idempotent matrix with rank $n - r$ and use the following theorem:

Theorem 3. Let X be a random vector with $E(X) = 0$ and $V(X) = \sigma^2 I$; and A be an idempotent vector with rank r then

$$E(X'AX) = r\,\sigma^2 . \tag{6}$$

Now ;

$$\begin{aligned} SSE &= (\beta' X' + \varepsilon')(I - XGX')(X\beta + \varepsilon) \\ &= \varepsilon' (I - XGX')\,\varepsilon . \end{aligned} \tag{7}$$

After using (6) in (7)

$$E(SSE) = (n - r)\,\sigma^2 , \quad (r = r(X))$$

and so an unbiased estimator of o^2 is

$$\widehat{\sigma}^2 = \frac{SSE}{n - r} .$$

7. ESTIMABLE FUNCTIONS

We saw that $\widehat{\beta}$ is not an unbiased estimator of β. But certain linear combinations of the elements of the solutions $\widehat{\beta}$ have a unique and unbiased value no matter what solution of β is used, these combinations are $Q'\beta$ where Q' is such that $Q'H = Q'$ and called estimable functions.

For any arbitrary vector W', $W'H\beta$ is an estimable function.

Since

$$Q' = W'H \quad \rightarrow \quad Q'H = W'H^2 = W'H = Q' .$$

Hence for any vector W', $W'H\beta$ is unique and $E(Q'\widehat{\beta}) = Q'\beta$. To show this:

$$\begin{aligned} Q'\widehat{\beta} &= W'H\widehat{\beta} \\ &= W'GX'Y \end{aligned}$$

and

$$\begin{aligned} E(Q'\widehat{\beta}) &= W'GX'E(Y) \\ &= W'GX'X\beta \\ &= Q'\beta. \end{aligned}$$

After making the necessary calculation, the variance of this estimable function and its estimator are found as

$$V(Q'\widehat{\beta}) = W'GW\sigma^2$$

$$\widehat{\sigma}^2 = \frac{(Y'Y - Y'XGX'Y)}{n-r} .$$

REFERENCES

[1] CARTER, W.H. and MYERS, R.H. : *Ortogonal Contrasts and the generalized inverse in fixed effects analysis of variance*, The Amer. Statis., **5** (1972), 32-34.

[2] MAZUNDAR, S., LIVE, C.C. and BRYCE, G.R. : *Correspondence Between a linear restriction and a generalized inverse in linear Model Analysis*, The Amer. Statis., 2 (1980), 103-104.

[3] MOORE, E.H. : *On the reciprocal of the general algebraic matrix* (Abstract). Bull. Amer. Math. Soc., 26 (1920), 394-395.

[4] MOORE, E.H. : *General Analysis*, American Philosophical Society, Philadelphia, 1935.

[5] PENROSE, R. : *A generalized inverse of matrices*, Proc. Camb. Phil. Soc., **51** (1955), 406-413.

[6] SEARLE, S.R. : *Restrictions and generalized inverses in linear models*, The Amer. Statis., **1** (1974), 53-54.

[7] SPEED, F.M. : *An application of the generalized inverse to the one way classification*, The Amer. Statis., 1 (1974), 16-18.

[8] RAO, C R. : Generalized Inverse of Matrices and its Applications, John Willey and Sons Inc., 1971.