

RECURSIVE ESTIMATION AND TESTING IN GENERAL LINEAR MODELS WITH APPLICATIONS TO REGRESSION DIAGNOSTICS

P. BHIMASANKARAM AND S. RAO JAMMALAMADAKA

1. Introduction

Recursive estimation of the Best Linear Unbiased Estimators (BLUES) and residual sum of squares in linear models has been dealt with, among others, by Plackett (1950), Mitra and Bhimasankaram (1971), McGillchrist and Sandland (1979), Haslett (1985) and Chib, Jammalamadaka and Tiwari (1987), when one or more additional observations become available. While the first two papers considered the model with uncorrelated observations the latter three papers deal with dependent errors. All papers except that of Mitra and Bhimasankaram consider full rank design matrix. Golub and Styan (1973), Paige (1978) and Kourouklis and Paige (1981) consider numerically stable recursive computations for the general linear model. However, the exact algebraic expressions for these recursions have not been available so far.

We presently consider a general linear model with a possibly rank deficient design matrix and a general positive definite dispersion matrix (corresponding to dependent observations). The exact algebraic expressions for recursive formulae for the BLUES as well as Likelihood Ratio Test (LRT) statistics are obtained in this paper for the addition or deletion of an observation or a parameter from the model. Surprisingly in several cases the LRT statistics remain unchanged although the least squares estimators are different. The recursive formulae for BLUES, Residual sum of squares and LRT statistics when an observation is deleted are very useful in regression diagnostics for detecting influential observations. The formulae we obtain in this context can be regarded as extensions of the formulae for DFBETA, DFBETAS, etc. of Belsley et. al. (1980) for general linear model. Addition and deletion of a parameter is a particularly important consideration in regression models as these correspond to the inclusion or exclusion of a predictor variable in the model.

Kourouklis and Paige (1981) give a nice computational method for recursive estimation and testing in a general linear model with a possibly singular dispersion matrix. However, the exact algebraic formulae for such recursions are not available and are cur-

rently under investigation.

We use the following notation. We consider only real matrices in this paper. For a matrix A , let $\rho(A)$, $\text{tr}(A)$, A' , $C(A)$ denote its rank, trace, transpose, and column space respectively. A^- denotes a generalized inverse (g -inverse) of A , i.e. A^- is a matrix satisfying $AA^-A = A$.

The ordered triplet (Y_h, X_h, Σ_h) denotes the linear model

$$Y_h = X_h\beta_h + \epsilon_h$$

where h is the number of observations and $E(\epsilon_h) = 0$ and $D(\epsilon_h) = \sigma^2\Sigma_h$. The $(h + 1)^{th}$ observation is denoted by $Y(h + 1)$ and Y_{h+1} denotes the complete observation vector with $h + 1$ components. The row vector in the design matrix X_{h+1} corresponding to the observation $Y(h + 1)$ is $X'(h + 1)$ and $\sigma^2\Sigma(h + 1)$ is the variance of $Y(h + 1)$. $R_{0_h}^2$ denotes the residual sum of squares, namely $(Y_h - X_h\hat{\beta}_h)'\Sigma_h^{-1}(Y_h - X_h\hat{\beta}_h)$ where $\hat{\beta}_h$ is any solution to the normal equations $(X_h'\Sigma_h^{-1}X_h)\beta_h = X_h'\Sigma_h^{-1}Y_h$.

Consider a hypothesis $H : A\beta_h = \xi$ where $C(A') \subseteq C(X')$.

Let $T_h = \text{Min}_{\beta_h: A\beta_h = \xi} (Y_h - X\beta_h)'\Sigma_h^{-1}(Y_h - X\beta_h)$. We use $R_{H_h}^2$ to denote $T_h - R_{0_h}^2$. Let $\hat{\beta}_h$ be as defined above. We use $D(A\hat{\beta}_h)$ to denote $A(X_h'\Sigma_h^{-1}X_h)^-A'$ for convenience. (Strictly speaking, $D(A\hat{\beta}_h) = \sigma^2 A(X_h'\Sigma_h^{-1}X_h)^-A'$.) Using Wald's representation we can write

$$R_{H_h}^2 = (A\hat{\beta}_h - \xi)'[D(A\hat{\beta}_h)]^-(A\hat{\beta}_h - \xi).$$

In section 2, we give some results on g -inverses of partitioned matrices and sums of matrices, some of which are new. In section 3, we give recursive formulae for BLUES and LRTS when an observation is added or deleted. Section 4 deals with recursive estimation and testing, when a parameter is added to or deleted from the model. In section 5, we give an application of these formulae to regression diagnostics.

2. Some results on the g -inverses of matrices

In this section, we give some results on g -inverses of a sum of matrices and of partitioned matrices which will be needed in the sequel.

Lemma 2.1. *Let A be a symmetric matrix of order $n \times n$ and let A^- be a g -inverse of A . Let u be a column vector with n components. Case(1). $u \notin C(A)$. Then for any non-zero real number k ,*

$$(A + kuu')^- = (I - bu')A^-(I - ub') + \frac{bb'}{k}$$

where $c = (I - A^{-1}A)(I - AA^{-1})u$ and $b = \frac{c}{c'u}$. Case(2). $u \in C(A)$. Then for any non-zero real number k ,

$$(A + kuu')^- = \begin{cases} A^- & \text{if } u'A^-u = \frac{1}{k} \\ A^- - \frac{A^-uu'A^-}{\frac{1}{k} + u'A^-u} & \text{otherwise.} \end{cases}$$

Proof of Case (1) follows, once it is observed that $b'u = 1$ and $Ab = 0$. Case (2) is well-known for $k = 1$ and $u'Au^- \neq -1$. (See for example Rao and Mitra (1971), pp. 40). The other part can be directly verified.

Lemma 2.2. (Rohde (1965), Bhimasankaram (1971)). Let Σ be a nonnegative definite (nnd) matrix. Partition $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ where Σ_{11} and Σ_{22} are square matrices. Let Σ_{11}^- be a g-inverse of Σ_{11} . Denote $F = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^-\Sigma_{12}$. Then one choice of a g-inverse of Σ is

$$\begin{bmatrix} \Sigma_{11}^- & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Sigma_{11}^-\Sigma_{12} \\ -I \end{bmatrix} F^-(\Sigma_{21}\Sigma_{11}^- - I)$$

where F^- is any g-inverse of F .

Lemma 2.3. (Bhimasankaram (1988)). Let $M = \begin{bmatrix} A & x \\ x' & c \end{bmatrix}$ be an nnd matrix of order $(n + 1) \times (n + 1)$ and let A be of order $n \times n$. Let $G = \begin{bmatrix} B & y \\ y' & d \end{bmatrix}$ be an nnd g-inverse of M and let B be of order $n \times n$. Then a choice of nnd g-inverse of A is

$$A^- = \begin{cases} (I_{\theta}^{\frac{1}{2}}(Bx + cy))G(I_{\theta}^{\frac{1}{2}}(Bx + cy))' & \text{if } x'y + cd \neq 1 \\ B - \frac{1}{d}yy' & \text{if } x'y + cd = 1, d \neq 0 \text{ and } Ay + dx = 0 \\ (I\xi)G(I\xi)' & \text{if } x'y + cd = 1, d \neq 0 \text{ and } Ay + dx \neq 0 \end{cases}$$

where $\theta = 1 - x'y - cd$ and $\xi = \|(Ay + dx)\|^{-2} \cdot (I - BA - yx') \cdot (Ay + dx)$.

Lemma 2.4. Consider the same set up as in Lemma 2.3. Then $\rho(M) = \rho(A) + 1$ if and only if $x'y + cd = 1, d \neq 0$ and $Ay + dx = 0$.

The proof follows by examining the difference between $\text{tr } MG$ and $\text{tr } AA^-$ where A^- is as in Lemma 2.3.

3. Updating formulae when an observation is introduced or deleted

In this section we obtain correction terms to the estimators of β and σ^2 in the linear model when an additional observation is introduced into the model or an observation is deleted from the model. We also obtain correction term for the likelihood ratio test for testing $H : A\beta = \xi$ in the above situations.

3(a). Introducing an additional observation

Consider the model (Y_h, X_h, Σ_h) and assume the following have been computed and are available to us:

$$\Sigma_h^{-1}; (X_h'\Sigma_h^{-1}X_h)^-; \hat{\beta}_h = (X_h'\Sigma_h^{-1}X_h)^-X_h'\Sigma_h^{-1}Y_h$$

$$\begin{aligned}
 R_{0h}^2 &= Y_h' X_h' \Sigma_h^{-1} Y_h - Y_h' \Sigma_h^{-1} X_h \hat{\beta}_h, \\
 [D(A\hat{\beta}_h)]^- &= [A(X_h' \Sigma_h^{-1} X_h)^- A']^- \quad \text{and} \\
 R_{H_h}^2 &= (A\hat{\beta}_h - \xi)' [D(A\hat{\beta}_h)]^- (A\hat{\beta}_h - \xi)
 \end{aligned}
 \tag{3.1}$$

where $A\beta$ is estimable. We want to incorporate one more observation $Y(h + 1)$ with

$$\begin{aligned}
 E(Y(h + 1)) &= X'(h + 1)\beta, \\
 V(Y(h + 1)) &= \Sigma(h + 1)\sigma^2, \\
 \text{and } cov(Y(h + 1), Y_h) &= c_h' \sigma^2
 \end{aligned}$$

into the analysis and update the formulae for the estimators and test statistics. Denoting

$$Y_{h+1} = \begin{bmatrix} Y_h \\ Y_{(h+1)} \end{bmatrix}, \quad X_{h+1} = \begin{bmatrix} X_h \\ X'(h + 1) \end{bmatrix} \quad \text{and} \quad \Sigma_{h+1} = \begin{bmatrix} \Sigma_h & C_h \\ C_h' & \Sigma(h + 1) \end{bmatrix},$$

we can represent the new model by $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$.

First we make the following computations using the available information (3.1).

$$\begin{aligned}
 \alpha &= \Sigma(h + 1) - c_h' \Sigma_h^{-1} c_h. \\
 v &= X_h' \Sigma_h^{-1} c_h - X(h + 1). \\
 d &= c_h' \Sigma_h^{-1} Y_h - Y(h + 1) \\
 w &= (X_h' \Sigma_h^{-1} X_h)^- v \\
 \theta &= v' w \\
 B &= I - (X_h' \Sigma_h^{-1} X_h)^- (X_h' \Sigma_h^{-1} X_h)
 \end{aligned}$$

and $b = \frac{\alpha^{1/2}}{X(h+1)' B B' X(h+1)} \cdot B B' X(h + 1)$ if $B' X(h + 1) \neq 0$

We consider two cases, namely, $X_h(h + 1) \in C(X_h')$ and $X_h(h + 1) \notin C(X_h')$. Notice that $X_{(h+1)} \in C(X_h')$ if and only if $B' X(h + 1) = 0$.

We prove

Theorem 3.1. *Let $B' X(h + 1) = 0$. Then the following hold.*

- (a) $\hat{\beta}_{h+1} = \hat{\beta}_h + \frac{d-v'\hat{\beta}_h w}{\alpha+\theta}$
- (b) $R_{0_{h+1}}^2 = R_{0_h}^2 + \frac{(d-v'\hat{\beta}_h)^2}{\alpha+\theta}$ and
- (c)

$$R_{H_{h+1}}^2 = \begin{cases} R_{H_h}^2 + \frac{(d-v'\hat{\beta}_h)}{\alpha+\theta} (2\mu + d - v'\hat{\beta}_h) & \text{if } \delta = \alpha + \theta \\ R_{H_h}^2 + \frac{(d-v'\hat{\beta}_h)}{\alpha+\theta} (2\mu + \frac{\delta}{\alpha+\theta} (d - v'\hat{\beta}_h)) \\ + \frac{1}{\alpha+\theta-\delta} \left[\mu + \frac{\delta(d-v'\hat{\beta}_h)}{(\alpha+\theta)} \right]^2 & \text{otherwise} \end{cases}$$

where $\delta = w' A' (A(X_h' \Sigma_h^{-1} X_h)^- A')^- A w$ and $\mu = w' A' [A(X_h' \Sigma_h^{-1} X_h)^- A']^- (A\hat{\beta}_h - \xi)$

Proof. Since $B'X(h + 1) = 0$, we have $X(h + 1) \in C(X'_h \Sigma_h^{-1} X_h) = C(X'_h)$. So $v = X'_h u - X(h + 1) \in C(X'_h \Sigma_h^{-1} X_h)$. It is easy to verify that $X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1} = X'_h \Sigma_h^{-1} X_h + \frac{1}{\alpha} v v'$ and $X'_{h+1} \Sigma_{h+1}^{-1} Y_{h+1} = X'_h \Sigma_h^{-1} Y_h + \frac{d}{\alpha} v$ in view of Lemma 2.2. Since $v \in C(X'_h \Sigma_h^{-1} X_h)$, $\theta = v'(X'_h \Sigma_h^{-1} X_h)^- v$ is nonnegative and since Σ_{h+1} is positive definite, $\alpha > 0$. So, $1 + \frac{\theta}{\alpha} \neq 0$ and by Lemma 2.1, Case (2), we have

$$\begin{aligned} (X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- &= (X'_h \Sigma_h^{-1} X_h + \frac{1}{\alpha} v v')^- \\ &= (X'_h \Sigma_h^{-1} X_h)^- - \frac{1}{\alpha + \theta} w w'. \end{aligned} \tag{3.2}$$

Now, $\hat{\beta}_{h+1} = (X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- X'_{h+1} \Sigma_{h+1}^{-1} Y_{h+1}$ which easily simplifies to the expression given in (a). This completes the proof of (a). To prove (b), notice that ,

$$\begin{aligned} R^2_{0_{h+1}} &= Y'_{h+1} \Sigma_{h+1}^{-1} (Y_{h+1} - X_{h+1} \hat{\beta}_{h+1}) \\ &= Y'_h \Sigma_h^{-1} Y_h + \frac{d^2}{\alpha} - Y'_h \Sigma_h^{-1} X_h \hat{\beta}_{h+1} - \frac{d}{\alpha} v' \hat{\beta}_{h+1} \end{aligned}$$

which simplifies to the expression given in (b) by substituting the value of $\hat{\beta}_{h+1}$ as obtained in (a). This completes the proof of (b). To prove (c), recall that

$$R^2_{H_{h+1}} = (A \hat{\beta}_{h+1} - \xi)' [A(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- A']^- (A \hat{\beta}_{h+1} - \xi). \tag{3.3}$$

Now, using (3.2) we have,

$$A(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- A' = A(X'_h \Sigma_h^{-1} X_h)^- A' - \frac{1}{\alpha + \theta} A w w' A'$$

From Lemma 2.1 Case (2), it follows that

$$\begin{aligned} &[A(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- A']^- \\ &= \begin{cases} [A(X'_h \Sigma_h^{-1} X_h)^- A']^- & \text{if } \delta = \alpha + \theta \\ [A(X'_h \Sigma_h^{-1} X_h)^- A']^- \\ + \frac{1}{\alpha + \theta - \delta} [A(X'_h \Sigma_h^{-1} X_h)^- A']^- A w w' A' [A(X'_h \Sigma_h^{-1} X_h)^- A']^- & \text{otherwise} \end{cases} \end{aligned} \tag{3.4}$$

Further,

$$A \hat{\beta}_{h+1} - \xi = A \hat{\beta}_h - \xi + \frac{d - v' \hat{\beta}_h}{\alpha + \theta} \cdot A w. \tag{3.5}$$

Now, (c) follows by substituting (3.4) and (3.5) in (3.3) and simplifying the expressions.

Remark. The LRT statistic having F -distribution under $H : AB = \xi$ in this case is

$$\frac{R^2_{H_{h+1}} / \rho(A)}{R^2_{0_{h+1}} / (h + 1 - \rho(x_h))}$$

since $\rho(X_{h+1}) = \rho(X_h)$.

We shall now consider the case $X(h+1) \notin C(X'_h)$ and prove the following somewhat unexpected result.

Theorem 3.2. *Let $B'X(h+1) \neq 0$. Then the following hold.*

- (a) $\hat{\beta}_{h+1} = \hat{\beta}_h + (d - v'\hat{\beta}_h)b$
- (b) $R^2_{0_{h+1}} = R^2_{0_h}$
- (c) $R^2_{H_{h+1}} = R^2_{H_h}$

Proof. Recall that $X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1} = X'_h\Sigma^{-1}_hX_h + \frac{1}{\alpha}vv'$ and $X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1}^{-1}Y_{h+1} = X'_h\Sigma^{-1}_hY_h + \frac{d}{\alpha}v$. Now $B'X(h+1) \neq 0$ if and only if $v \notin C(X'_h\Sigma^{-1}_hX_h)$. So, using Lemma 2.1 to get a g -inverse of $(X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1})$ and substituting it in the formula for $\hat{\beta}_{h+1}$ (see (3.1)), it can easily be shown that $\hat{\beta}_{h+1} = \hat{\beta}_h + (d - v'\hat{\beta}_h)b$. This proves (a). To prove (b), first notice that $X_hb = 0$ and $v'b = 1$. Now,

$$\begin{aligned} R^2_{0_{h+1}} &= Y'_{h+1}\Sigma^{-1}_{h+1}Y_{h+1} - Y'_{h+1}\Sigma^{-1}_{h+1}X_{h+1}\hat{\beta}_{h+1} \\ &= R^2_{0_h} + \frac{d^2}{\alpha} - \frac{d}{\alpha}(v'\hat{\beta}_h + d - v'\hat{\beta}_h) \\ &= R^2_{0_h}. \end{aligned}$$

This completes the proof of (b). To prove (c), observe that $C(A') \subseteq C(X')$ and hence $Ab = 0$. So, $[A(X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1})^{-1}A']^{-1} = (A(X'_h\Sigma^{-1}_hX_h)^{-1}A')^{-1}$ and $A\hat{\beta}_{h+1} = A\hat{\beta}_h$. Now (c) follows immediately.

This leads us to the following surprising

Corollary 3.3. *Let $B'X(h+1) \neq 0$. Let $C(A') \subseteq C(X'_h)$ and let $\xi \in C(A)$. Then the value of LRT statistic for testing $A\beta = \xi$ is the same under both the models (Y_h, X_h, Σ_h) and $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$.*

Proof. Since $B'X(h+1) \neq 0$, we have $\rho(X_{h+1}) = \rho(X_h) + 1$. Now, the LRT statistic for testing $A\beta = \xi$ under $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ is

$$\frac{R^2_{H_{h+1}}/\rho(A)}{R^2_{0_{h+1}}/(h+1-\rho(x_{h+1}))} = \frac{R^2_{H_h}/\rho(A)}{R^2_{0_h}/(h-\rho(x_h))}$$

which is precisely the LRT statistic for testing $A\beta = \xi$ under (Y_h, X_h, Σ_h) .

3(b). Deleting an observation

Here we consider the model $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ and assume that the following have computed and are available to us:

$$\Sigma^{-1}_{h+1}; (X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1})^{-1}; \quad \hat{\beta}_{h+1} = (X'_{h+1}\Sigma^{-1}_{h+1}X_{h+1})^{-1}X'_{h+1}\Sigma^{-1}_{h+1}Y_{h+1}$$

$$R_{0_{h+1}}^2 = Y'_{h+1} \Sigma_{h+1}^{-1} Y_{h+1} - Y'_{h+1} \Sigma_{h+1}^{-1} X_{h+1} \hat{\beta}_{h+1};$$

$$[D(A\hat{\beta}_{h+1})]^- = [A(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- A']^-$$

and

$$R_{H_{h+1}}^2 = (A\hat{\beta}_{h+1} - \xi)' [D(A\hat{\beta}_{h+1})]^- (A\hat{\beta}_{h+1} - \xi) \tag{3.8}$$

where $A\beta$ is estimable.

Now, consider the model (Y_h, X_h, Σ_h) obtained from $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ by deleting the last observation.

Write $\Sigma_{h+1}^{-1} = \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{12'} & \Sigma^{22} \end{bmatrix}$ where $(\Sigma^{12'} \Sigma^{22})$ is the last row of Σ_{h+1}^{-1} . It is easy to check that

$$\Sigma_h^{-1} = \Sigma^{11} - \frac{1}{\Sigma^{22}} \Sigma^{12} \Sigma^{12'}$$

and hence

$$\Sigma_{h+1}^{-1} = \begin{bmatrix} \Sigma_h^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\Sigma^{22}} \begin{bmatrix} \Sigma^{12} \\ \Sigma^{22} \end{bmatrix} (\Sigma^{12'} \Sigma^{22}).$$

It now follows that

$$X'_h \Sigma_h^{-1} X_h = X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1} - \frac{1}{\Sigma^{22}} uu' \tag{3.7}$$

where

$$u = X'_h \Sigma^{12} + \Sigma^{22} X(h+1). \tag{3.8}$$

Clearly, $u \in C(X'_{h+1}) = C(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})$. So, in view of Lemma 2.1 Case (2), we have

$$(X'_h \Sigma_h^{-1} X_h)^- = \begin{cases} (X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- & \text{if } u'v = \Sigma^{22} \\ (X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- + \frac{vv'}{\Sigma^{22} - u'v} & \text{if } u'v \neq \Sigma^{22} \end{cases} \tag{3.9}$$

where

$$v = (X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- u \tag{3.10}$$

Further,

$$X'_h \Sigma_h^{-1} Y_h = X'_{h+1} \Sigma_{h+1}^{-1} Y_{h+1} - \frac{d}{\Sigma^{22}} u \tag{3.11}$$

where

$$d = \Sigma^{12'} Y_h + \Sigma^{22} Y(h+1). \tag{3.12}$$

We first prove

Lemma 3.4. $X(h+1) \notin C(X'_h)$ if and only if $\Sigma^{22} = u'v$.

Proof. To prove the if part, let $\Sigma^{22} = u'v$.

Then

$$\begin{aligned} \rho(X_h) &= \rho(X'_h \Sigma_h^{-1} X_h) = \text{tr}[(X'_h \Sigma_h^{-1} X_h)(X'_h \Sigma_h^{-1} X_h)^-] \\ &= \text{tr}[(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1} - \frac{1}{\Sigma^{22}} uu')(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^-] \\ &= \rho(X_{h+1}) - \frac{u'v}{\Sigma^{22}} = \rho(X_{h+1}) - 1. \end{aligned}$$

Hence $X(h + 1) \notin C(X'_h)$. To prove the only if part, let $\Sigma^{22} \neq u'v$. Then,

$$\begin{aligned} \rho(X_h) &= \text{tr}[(X'_h \Sigma'_h X_h)(X'_h \Sigma_h^{-1} X_h)^-] \\ &= \text{tr}[(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1} + \frac{1}{\Sigma^{22}} uu') \left[(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- + \frac{vv'}{\Sigma^{22} - u'v} \right)] \\ &= \rho(X_{h+1}). \end{aligned}$$

Hence $X(h + 1) \in C(X'_h)$. We now prove

Theorem 3.5. Consider the models $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ and (Y_h, X_h, Σ_h) and let u, v and d be as defined in (3.8), (3.10) and (3.12) respectively. Let $X(h + 1) \notin C(X'_h)$. Then the following hold:

- (a) $\hat{\beta}_h = \hat{\beta}_{h+1} - \frac{d}{\Sigma^{22}} v$.
- (b) Let $A\beta$ be estimable under $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$. Then $A\beta$ is estimable under (Y_h, X_h, Σ_h) if and only if $Av = 0$.
- (c) $R_{0_h}^2 = R_{0_{h+1}}^2$
- (d) Let $A\beta$ be estimable under (Y_h, X_h, Σ_h) . Then $R_{H_h}^2 = R_{H_{h+1}}^2$

Proof. In view of Lemma 3.4, (a) follows easily from (3.9) and (3.11). To prove (b), let $C(A') \subseteq C(X'_{h+1})$. Now observe that $A\beta$ is estimable under (Y_h, X_h, Σ_h) if and only if

$$(X'_h \Sigma_h^{-1} X_h)(X'_h \Sigma_h^{-1} X_h)^- A' = A'. \tag{3.13}$$

From (3.7) and (3.9) it now follows that (3.13) holds if and only if $A(X'_{h+1} \Sigma_{h+1}^{-1} X_{h+1})^- uu' = 0$, or equivalently, if and only if $Av = 0$. To prove (c), first show that $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, the last column of the identity matrix, belongs to $C(X_{h+1})$. For this, let $\ell' X_n + m \cdot X'(h + 1) = 0$. This clearly means that $m = 0$ since $X(h + 1) \notin C(X'_h)$. So $\ell' 0 + m \cdot 1 = 0$. Hence $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in C(X_{h+1})$. Now, it immediately follows that $u' \hat{\beta}_{h+1} = d$. Now (c) follows by straightforward computation while (d) is verified easily using (3.9).

Remark 1. Let $X(h + 1) \notin C(X'_h)$ and let $p'\beta$ be any estimable function under (Y_h, X_h, Σ_h) . Then $p'\hat{\beta}_h = p'\hat{\beta}_{h+1}$ since $p'v = 0$ in view of (b) of the preceding theorem. So one can choose $\hat{\beta}_h = \hat{\beta}_{h+1}$ without loss of generality.

Remark 2. Lemma 3.4 and theorem 3.5 (c) and (d) also lead to Corollary 3.3.

Theorem 3.6. Consider the models $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ and (Y_h, X_h, Σ_h) and let u, v and d be as defined in (3.8), (3.10) and (3.12) respectively. Let $X(h + 1) \in C(X'_h)$. Then the following hold:

- (a) $\hat{\beta}_h = \hat{\beta}_{h+1} + \frac{(u' \hat{\beta}_{h+1} - d)v}{\Sigma^{22} - u'v}$,
- (b) $R_{0_h}^2 = R_{0_{h+1}}^2 + \frac{(u' \hat{\beta}_{h+1} - d)^2}{\Sigma^{22} - u'v}$,
- (c) If $A\beta$ is estimable under $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$, then it is also estimable under

(Y_h, X_h, Σ_h) and

$$R_{H_h}^2 = \begin{cases} R_{H_{h+1}}^2 - \frac{(u'\hat{\beta}_{h+1}-d)}{\Sigma^{22}-u'v} \{u'\hat{\beta}_{h+1} - d - 2w'(A\hat{\beta}_{h+1} - \xi)\} & \text{if } c = u'v - \Sigma^{22} \\ R_{H_{h+1}}^2 - \frac{(u'\hat{\beta}_{h+1}-d)^2}{\Sigma^{22}-u'v} - \frac{(u'\hat{\beta}_{h+1}-d+w'(A\hat{\beta}_{h+1}-\xi))^2}{\Sigma^{22}-u'v+c} & \text{otherwise} \end{cases}$$

where $w = [A(X'_{h+1}\Sigma_{h+1}^{-1}X_{h+1})^{-1}A']^{-1}Av$ and $c = v'A'w$.

Proof follows by straightforward verification using Lemma 2.1, Case (2).

4. Updating formulae when a parameter is introduced or deleted

In this section we obtain correction terms for the estimators of β and σ^2 when an additional parameter is introduced into the model or a parameter is deleted from the model. We also obtain correction terms for the likelihood ratio testing $A\beta = \xi$ in the above situations.

(a) Introducing an additional parameter.

Consider the model (Y, X, Σ) and assume the following have been computed and are available to us:

$$\Sigma^{-1}; \quad (X'\Sigma^{-1}X)^{-}; \quad \hat{\beta} = (X'\Sigma^{-1}X)^{-}X'\Sigma^{-1}Y;$$

$$R_0^2 = Y'\Sigma^{-1}Y - Y'\Sigma^{-1}X\hat{\beta};$$

$$[D(A\hat{\beta})]^{-} = (A(X'\Sigma^{-1}X)^{-}A')^{-}$$

and

$$R_H^2 = (A\hat{\beta} - \xi)'[D(A\hat{\beta})]^{-}(A\hat{\beta} - \xi) \text{ where } A\beta \text{ is estimable.}$$

We want to incorporate one more parameter ν into the model and this introduces a new column x into the design matrix thus the new model can be written as

$$Y = (X : x)\theta + \epsilon \tag{4.1}$$

where $\theta = \begin{pmatrix} \beta \\ \nu \end{pmatrix}$, $E(\epsilon) = 0$ and $D(\epsilon) = \Sigma\sigma^2$.

Writing $W = (X : x)$ and $\theta = \begin{pmatrix} \beta \\ \nu \end{pmatrix}$ we can denote the model by (Y, W, Σ) where the parametric vector is θ . We are interested in computing $\hat{\theta}$, the usual unbiased estimator of θ and likelihood ratio test statistic for $A\beta = \xi$ if $A\beta$ is still estimable. We also want to test whether it is worthwhile to introduce the new parameter ν .

For convenience, we refer to (Y, X, Σ) as the old model and (Y, W, Σ) as the new model. Accordingly R_0^2 (old), R_0^2 (new), R_H^2 (old), R_H^2 (new) are defined as R_0^2 and R_H^2 respectively under old and new models. Now,

$$W'\Sigma^{-1}W = \begin{bmatrix} X'\Sigma^{-1}X & X'\Sigma^{-1}x \\ x'\Sigma^{-1}X & x'\Sigma^{-1}x \end{bmatrix}$$

In view of Lemma 2.2, it follows that

$$(W'\Sigma^{-1}W)^{-} = \begin{cases} \begin{bmatrix} (X'\Sigma^{-1}X)^{-} & 0 \\ 0 & 0 \end{bmatrix} & \text{if } \sigma = 0 \\ \begin{bmatrix} (X'\Sigma^{-1}X)^{-} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\alpha} \begin{pmatrix} u \\ -1 \end{pmatrix}^{(u'-1)} & \text{otherwise} \end{cases} \quad (4.2)$$

where

$$u = (X'\Sigma^{-1}X)^{-}X'\Sigma^{-1}x. \quad (4.3)$$

and

$$\alpha = x'\Sigma^{-1}x - x'\Sigma^{-1}Xu \quad (4.4)$$

It is easy to check that $\alpha = 0$ if and only if $x \in C(X)$. The following results can be proved by straightforward verification.

Theorem 4.1. *Consider the model (Y, W, Σ) as in (4.1), and let $x \in C(X)$. Then the following hold:*

- (a) $\tilde{\theta} = \begin{pmatrix} \hat{\beta} \\ 0 \end{pmatrix}$,
- (b) $R_{0(\text{new})}^2 = R_{0(\text{old})}^2$,
- (c) Let $A\beta$ be estimable under the old model. Then it is estimable under the new model iff $Au = 0$,
- (d) Let $Au = 0$. Then $R_{H(\text{new})}^2 = R_{H(\text{old})}^2$,
- (e) ν is not estimable under the new model.

Corollary 4.2. *Under the set up of theorem 4.1, let $A\beta$ be estimable under the old model as well as the new model (i.e., $AX^{-}x = 0$). Then the LRT statistics for testing $A\beta = \xi$ in the old and new models are the same and have F distribution with the same degrees of freedom under the null hypothesis.*

Theorem 4.3. *Consider the model (Y, W, Σ) as in (4.1) and let $x \notin C(X)$. Then the following hold:*

- (a) $\tilde{\theta} = \begin{pmatrix} \tilde{\beta} \\ \tilde{\nu} \end{pmatrix}$ where $\tilde{\nu} = \frac{1}{\alpha}x'\Sigma^{-1}(Y - X\hat{\beta})$ and $\tilde{\beta} = \hat{\beta} - \tilde{\nu}u$.
- (b) $R_{0(\text{new})}^2 = R_{0(\text{old})}^2 + \tilde{\nu} \cdot Y'\Sigma^{-1}(Xu - x)$.
- (c) Let $A\beta$ be estimable under the old model. Then $A\beta$ is estimable under the new model and

$$R_{H(\text{new})}^2 = \begin{cases} R_{H(\text{old})}^2 - \tilde{\nu}^2\alpha - 2\tilde{\nu}v'(A\hat{\beta} - \xi) & \text{if } v'Au = -\alpha \\ R_{H(\text{old})}^2 + \tilde{\nu}^2v'Au - 2\tilde{\nu}v'(A\hat{\beta} - \xi) - \frac{(v'(A\hat{\beta} - \xi))^2}{\alpha + v'Au} & \text{otherwise} \end{cases}$$

- (d) ν is estimable under (Y, W, Σ) and to test $H : \nu = 0$, the usual statistic is

$$\frac{x'\Sigma^{-1}(Y - X\hat{\beta})}{\sqrt{(x - Xu)'\Sigma^{-1}(x - Xu) \cdot \frac{R_{0(\text{new})}^2}{n - \rho(x) - 1}}}$$

which has the student's t distribution with $n - \rho(x) - 1$ degrees of freedom under H .

Again, the proof is computational and is omitted.

(B) Deleting a parameter

Consider the model (Y, X, Σ) and assume the following have been computed and are available to us

$$\Sigma^{-1}; \quad (X'\Sigma^{-1}X)^{-}; \quad X'\Sigma^{-1}Y; \quad \hat{\beta} = (X'\Sigma^{-1}X)^{-}X'\Sigma^{-1}Y;$$

$$R_0^2 = Y'\Sigma^{-1}Y - Y'\Sigma^{-1}X\hat{\beta};$$

$$[D(A\hat{\beta})]^{-} = (A(X'\Sigma^{-1}X)^{-}A')^{-}$$

and $R_H^2 = (A\hat{\beta} - \xi)'[D(A\hat{\beta})]^{-}(A\hat{\beta} - \xi)$ where $A\beta$ is estimable. Further let $(X'\Sigma^{-1}X)^{-}$ be an nnd g -inverse of $X'\Sigma^{-1}X$. We want to delete the last component of the parametric vector β from the model and analyse the resulting model. Let us write

$$X = (W : x) \quad \text{and} \quad \beta' = (\tau' : \mu) \tag{4.5}$$

where x is the last column of X and μ is the last component of β . Then the resulting model is

$$Y = W\tau + \epsilon, \quad E(\epsilon) = 0, \quad D(\epsilon) = \sigma^2\Sigma. \tag{4.6}$$

Now

$$X'\Sigma^{-1}X = \begin{bmatrix} W'\Sigma^{-1}W & W'\Sigma^{-1}x \\ x'\Sigma^{-1}W & x'\Sigma^{-1}x \end{bmatrix} \tag{4.7}$$

and $X'\Sigma^{-1}Y = \begin{bmatrix} W'\Sigma^{-1}Y \\ x'\Sigma^{-1}Y \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, say. (4.8)

Partition the nnd g -inverse $(X'\Sigma^{-1}X)^{-}$ as

$$(X'\Sigma^{-1}X)^{-} = \begin{bmatrix} B & v \\ v' & c \end{bmatrix} \tag{4.9}$$

where $(v' c)$ is a row vector and c is a scalar. Notice that $x \in \mathbf{C}(W)$ if and only of

$$\rho(W'\Sigma^{-1}W) = \rho(W) = \rho(X) = \rho(X'\Sigma^{-1}X).$$

We consider the two cases, namely, $x \in \mathbf{C}(W)$ and $x \notin \mathbf{C}(W)$ in the following two theorems.

Case 1. Let $x \notin \mathbf{C}(W)$. Then from Lemma 2.4, we have $c \neq 0$, $W'\Sigma^{-1}(Wv + cx) = 0$ and $x'\Sigma^{-1}Wv + cx'\Sigma^{-1}x = 1$. The following theorem gives the new least squares estimator $\tilde{\tau}$, new R_0^2 and new R_H^2 when the hypothesis is testable.

Theorem 4.4. Consider the set up as in (4.5) and (4.6) and let $x \notin C(W)$. Then the following hold:

- (a) $\tilde{\tau} = \hat{\tau} - (u_2 + \frac{1}{c}v'u_1)v$;
- (b) $R_{0_{new}}^2 = R_{0_{old}}^2 + \frac{1}{c}(v'u_1 + cu_2)^2 - (1 - \frac{1}{c})u_2 \cdot v'u_1$;
- (c) Let $A\tau$ be estimable under both old and new models. Consider $H : A\tau = \xi$ against $H_a : A\tau \neq \xi$. Then

$$R_{H_{new}}^2 = \begin{cases} R_{H_{(old)}}^2 + c(u_2 + \frac{1}{c}v'u_1)^2 - 2(u_2 + \frac{1}{c}v'u_1) & \text{if } c = v'A'R^{-}Av \\ R_{H_{(old)}}^2 + v'A'R^{-}Av(u_2 + \frac{1}{c}v'u_1)^2 \\ -2(u_2 + \frac{1}{c}v'u_1)p - \frac{p^2}{v'A'R^{-}Av - c} & \text{otherwise} \end{cases}$$

where $R = A(X'\Sigma^{-1}X)^{-1}A'$, u_1, u_2, v and c are as defined in (4.8) and (4.9) and $p = v'A'R^{-}(A\hat{\tau} - \xi)$.

Proof is computational and we omit.

Case 2: Let $x \in C(W)$. Then, from Theorem 4.1, we can establish

Theorem 4.5. Consider the set up as in (4.5) and (4.6) and let $x \in C(W)$. Then the following hold:

- (a) For every parametric function $p'\tau$ estimable under both the models, (Y, X, Σ) and (Y, W, Σ) ,

$$p'\tilde{\tau} = p'\hat{\tau}.$$

- (b) $R_{0_{(new)}}^2 = R_{0_{(old)}}^2$
- (c) If $A\tau$ is estimable under both the models, then for testing $H : A\tau = \xi$ against $A\tau \neq \xi$, $R_{H_{(new)}}^2 = R_{H_{(old)}}^2$.

5. Application to Regression Diagnostics

Following Belsley, Kuh and Welsch (1980), in this section, we study the influence of an observation on the BLUE of an estimable parametric function, on the residual sum of squares, on the scaled measure of change and on the change of fit. Such a study is very important in regression diagnostics as a way of detecting influential observations as well as outliers. (See Belsley et. al. (1980)).

We consider the model $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ and study the effect of deleting the $(h + 1)$ st observation. The effect of deleting any other observation can be studied the same way by using suitable permutation of the rows/columns of the matrices and vectors involved.

We shall consider two cases, namely, $X(h + 1) \in C(X'_h)$ and $X(h + 1) \notin C(X'_h)$.

Case 1: $X(h + 1) \in C(X'_h)$. In this case, it is easily seen that any $p'\beta$ which is estimable under $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$, is also estimable under (Y_h, X_h, Σ_h) .

We now state the results in the form of a theorem, whose proof follows from Theorem 3.6.

Theorem 5.1. Consider the model $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$ and let $X(h+1) \in C(X'_h)$. Then

- a) The change in the BLUE of $p'\beta$ after deleting $(h+1)$ st observation, DF $BETA_{h+1,p}$ is given by

$$\begin{aligned} DF \quad BETA_{h+1,p} &= p'\hat{\beta}_{h+1} - p'\hat{\beta}_h \\ &= \frac{d - u'\hat{\beta}_{h+1}}{\Sigma^{22} - u'v} p'v \end{aligned}$$

- a') The square of the maximum possible change in an estimable parametric function $p'\beta$ such that $p'p = 1$ after deleting $(h+1)$ th at most observation is

$$DF \quad BETA_{h+1,MAX} = \frac{(d - u'\hat{\beta}_{h+1})^2}{(\Sigma^{22} - u'v)^2} \cdot v'v.$$

- b) The change in the residual sum of squares after deleting the $(h+1)$ th observation is

$$R_{0h}^2 - R_{0h+1}^2 = \frac{(u'\hat{\beta}_{h+1} - d)^2}{\Sigma_{22} - u'v}$$

- c) The change of fit of the $h+1$ th observation after deleting it is given by

$$\begin{aligned} DF \quad FIT_{h+1} &= \hat{y}_{h+1} - \hat{y}_{(h+1)} \\ &= \frac{(d - u'\hat{\beta}_{h+1})}{\Sigma^{22} - u'v} \cdot X'_{h+1}v. \end{aligned}$$

- d) The scaled measure of change in the BLUE of an estimable $p'\beta$ due to the deletion of $(h+1)$ st observation is given by

$$DF \quad BETAS_{h+1,p} = \frac{d - u'\hat{\beta}_{h+1}}{\Sigma^{22} - u'v} \cdot \frac{p'v}{\sqrt{\frac{R_{\sigma h}^2}{n-r} \cdot p'(X'_{h+1}\Sigma_{h+1}^{-1}X_{h+1})^{-p}}}$$

- e) The square of the maximum scaled measure of charge in the BLUE of any estimable parametric function due to the deletion of $(h+1)$ st observation is given by

$$DF \quad BETAS_{h+1,MAX} = \frac{(d - u'\hat{\beta}_{h+1})^2}{(\Sigma^{22} - u'v)^2} \cdot v'(x'_{h+1}\Sigma_{h+1}^{-1}x_{h+1})v$$

Case 2: $X(h+1) \notin C(X'_h)$.

Let $p'\beta$ be estimable under $(Y_{h+1}, X_{h+1}, \Sigma_{h+1})$. Then it is estimable under (Y_h, X_h, Σ_h) if and only if $p'v = 0$. (This follows from proof of (b) of theorem 3.5). Hence $p'\hat{\beta}_h = p'\hat{\beta}_{h+1}$ for all estimable $p'\beta$ under (Y_h, X_h, Σ_h) . So $(h+1)$ st observation does not have any influence on the BLUES of the parametric functions, estimable both the models.

References

- [1] D. A. Belsley, E. Kuh and R. E. Welsch, *Regression Diagnostics*, Wiley, New York, 1980.
- [2] P. Bhimasankaram, "On generalized inverse of partitioned matrices," *Sankhya, Ser. A*, 33(1971), 311-314.
- [3] P. Bhimasankaram, "On generalized inverse of a block in a partitioned matrix," to appear in *Linear Algebra and its Applications*, (1988).
- [4] S. Chib, S. R. Jammalamadaka, and R. C. Tiwari, "Another look at some results on the recursive estimation in the general linear model," *American Statistician*, 41(1987), 56-58.
- [5] G. H. Golub, and G. P. H. Styan, "Numerical computations for univariate linear models," *Journal of Statistical Computing and Simulation*, 2(1973), 253-274.
- [6] S. Haslett, "Recursive estimation of the general linear model with dependent errors multiple additional observations," *Australian Journal of Statistics*, 27(1985), 183-188.
- [7] S. Kourouklis, and C. C. Paige, "A constrained least squares approach to the general Gauss-Markov linear model," *Journal of the American Statistical Association*, 76(1981), 620-625.
- [8] C. A. McGillchrist, and R. L. Sandland, "Recursive estimation of the general linear model with dependent errors," *Journal of the Royal Statistical Society, Ser. B*, 41(1979), 65-68.
- [9] S. K. Mitra, and P. Bhimasankaram, "Generalized inverses of partitioned matrices and recalculation of least squares estimators for data or model changes," *Sankhya, Ser. A*, 33(1971), 396-410.
- [10] C. C. Paige, "Numerically stable computations for general univariate linear models," *Communications in Statistics B7*(5)(1978), 437-453.
- [11] R. L. Plackett, "Some theorems in least squares," *Biometrika*, 37(1950), 149-157.
- [12] C. A. Rohde, "Generalized inverse of partitioned matrices," *SIAM Journal of Appl. Math.*, 18 (1965), 1033-1035.

Statistics and Applied Probability Program, University of California, Santa Barbara, CA 93106, U.S.A.