An Efficient Algorithm For Case-Deletion in Mixed Linear Models

Kang-Sup Lee  
*Professor, Division of Natural Science, Dankook University*  
*San8, HanNam-Dong, YongSan-Gu, Seoul, Korea*  
*leeks@dankook.ac.kr*

Jang-Taek Lee  
*Professor, Division of Natural Science, Dankook University*  
*San8, HanNam-Dong, YongSan-Gu, Seoul, Korea*  
*jtlee@dankook.ac.kr*

1 Introduction

The last 10 years have seen an increased emphasis on case-deletion diagnostics to the mixed linear models. Being essentially a computational approach, case-deletion can produce exceedingly complex formulas in mixed linear models. Christensen et al.(1992) have provided computational formulas to the inverse of the variance matrix of data set for single case-deletion, by taking the appropriate elements of the inverse of the covariance matrix of the full data set. Hurtado(1993) presented similar results. However, if we examine case-deletion results for estimates of the variance components like MLE, REMLE and MINQUE, these methods need the inversion of variance matrix instead of using the W-transformation.

In this paper, we presents a simple and efficient algorithm for case-deletion using W-matrix. This algorithm makes deletion diagnostics practically.

2 A New Algorithm Using W-Matrix

For ML, REML and MINQUE equations, the following W-matrix is needed at each iterative step.

\[
W = \begin{pmatrix}
Z'H^{-1}Z & Z'H^{-1}X & Z'H^{-1}y \\
X'H^{-1}Z & X'H^{-1}X & X'H^{-1}y \\
y'H^{-1}Z & y'H^{-1}X & y'H^{-1}y
\end{pmatrix}.
\]  

(1)
Lee and Kim (1988) has provided an efficient algorithm for W-matrix using the reformation which is exactly identical to W-matrix. The W-matrix can be reconstructed as the terms of balanced design matrices \(X_0\) and \(Z_0\) having one and only one observation in each cell, replications matrix \(D\), the vector of cell means \(\bar{y}\), and error sum of squares \(s^2\). The reformation the same matrix as W-matrix is as follows.

\[
W^* = \begin{pmatrix}
Z_0' M Z_0 & Z_0' M X_0 & Z_0' M \bar{y} \\
X_0' M Z_0 & X_0' M X_0 & X_0' M \bar{y} \\
\bar{y}' M Z_0 & \bar{y}' M X_0 & \bar{y}' M \bar{y} + s^2
\end{pmatrix}, \tag{2}
\]

where \(M = (I_n + DU_0U_0')^{-1}D\).

This method requires inversions of an \(n \times n\) matrix at each iteration, where \(n\) is the number of nonempty cells. Applying this algorithm to case-deletion case, it is important that the algorithm to calculate M matrix after some observations have been eliminated be as efficient as possible. The next theorem is basic to efficient multiple-case deletion with relate to same cell in mixed linear models. For the simplicity of notation, let \(P\) matrix be \(P = (I + DU_0U_0')^{-1}\). Also, let \(M^*\) and \(D^*\) be a \(M\) and \(D\) matrix in (2) after \(k\) observations from the original data set in same cell have been deleted, respectively.

**Theorem 1.** \(M^* = (I_n + tw_iu_i')PD^*\), where \(u_i'\) is the \(i\)-th row of matrix \(U_0U_0'\), \(v_i\) is the \(i\)-th column of matrix \(P\) and constant \(t\) is defined by \(t = k/(1 - ku_i'v_i)\).

The result of above theorem can be extended to multiple-case deletion in connection with many cells.

**References**

