

0.1 The CPJ Paper

0.1.1 Case-Deletion results for Variance components

? examines case deletion results for estimates of the variance components, proposing the use of one-step estimates of variance components for examining case influence. The method describes focuses on REML estimation, but can easily be adapted to ML or other methods.

This paper develops their global influences for the deletion of single observations in two steps: a one-step estimate for the REML (or ML) estimate of the variance components, and an ordinary case-deletion diagnostic for a weighted regression problem (conditional on the estimated covariance matrix) for fixed effects.

0.1.2 CPJ Notation

$$\mathbf{C} = \mathbf{H}^{-1} = \begin{bmatrix} c_{ii} & \mathbf{c}'_i \\ \mathbf{c}_i & \mathbf{C}_{[i]} \end{bmatrix}$$

? noted the following identity:

$$\mathbf{H}^{-1}_{[i]} = \mathbf{C}_{[i]} - \frac{1}{c_{ii}} \mathbf{c}_{[i]} \mathbf{c}'_{[i]}$$

? use the following as building blocks for case deletion statistics.

- \check{x}_i
- \check{z}_i
- $\check{z}_i j$
- \check{y}_i
- $p_i i$
- m_i

All of these terms are a function of a row (or column) of \mathbf{H} and $\mathbf{H}^{-1}_{[i]}$

0.2 The CPJ Paper

0.2.1 Case-Deletion results for Variance components

? examines case deletion results for estimates of the variance components, proposing the use of one-step estimates of variance components for examining case influence. The method describes focuses on REML estimation, but can easily be adapted to ML or other methods.

This paper develops their global influences for the deletion of single observations in two steps: a one-step estimate for the REML (or ML) estimate of the variance components, and an ordinary case-deletion diagnostic for a weighted regression problem (conditional on the estimated covariance matrix) for fixed effects.

0.2.2 CPJ Notation

$$\mathbf{C} = \mathbf{H}^{-1} = \begin{bmatrix} c_{ii} & \mathbf{c}'_i \\ \mathbf{c}_i & \mathbf{C}_{[i]} \end{bmatrix}$$

? noted the following identity:

$$\mathbf{H}^{-1}_{[i]} = \mathbf{C}_{[i]} - \frac{1}{c_{ii}} \mathbf{c}_{[i]} \mathbf{c}'_{[i]}$$

? use the following as building blocks for case deletion statistics.

- \check{x}_i
- \check{z}_i
- $\check{z}_i j$
- \check{y}_i
- $p_i i$
- m_i

All of these terms are a function of a row (or column) of \mathbf{H} and $\mathbf{H}^{-1}_{[i]}$

0.3 Matrix Notation for Case Deletion

0.3.1 Case deletion notation

For notational simplicity, $\mathbf{A}(i)$ denotes an $n \times m$ matrix \mathbf{A} with the i -th row removed, a_i denotes the i -th row of \mathbf{A} , and a_{ij} denotes the (i, j) -th element of \mathbf{A} .

0.3.2 Partitioning Matrices

Without loss of generality, matrices can be partitioned as if the i -th omitted observation is the first row; i.e. $i = 1$.

0.4 CPJ's Three Propositions

Proposition 1

$$\mathbf{V}^{-1} = \begin{bmatrix} \nu^{ii} & \lambda'_i \\ \lambda_i & \Lambda_{[i]} \end{bmatrix}$$

$$\mathbf{V}_{[i]}^{-1} = \Lambda_{[i]} - \frac{\lambda_i \lambda'_i}{\lambda_i}$$

0.4.1 Proposition 2

$$(i) \quad \mathbf{X}_{[i]}^T \mathbf{V}_{[i]}^{-1} \mathbf{X}_{[i]} = \mathbf{X}' \mathbf{V}^{-1} \mathbf{X}$$

$$(ii) \quad = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{Y})^{-1}$$

$$(iii) \quad \mathbf{X}_{[i]}^T \mathbf{V}_{[i]}^{-1} \mathbf{Y}_{[i]} = \mathbf{X}' \mathbf{V}^{-1} \mathbf{Y}$$

0.4.2 Proposition 3

This proposition is similar to the formula for the one-step Newtown Raphson estimate of the logistic regression coefficients given by Pregibon (1981) and discussed in Cook Weisberg.

0.5 The CPJ Paper

0.5.1 Case-Deletion results for Variance components

? examines case deletion results for estimates of the variance components, proposing the use of one-step estimates of variance components for examining case influence. The method describes focuses on REML estimation, but can easily be adapted to ML or other methods.

This paper develops their global influences for the deletion of single observations in two steps: a one-step estimate for the REML (or ML) estimate of the variance components, and an ordinary case-deletion diagnostic for a weighted regression problem (conditional on the estimated covariance matrix) for fixed effects.

0.5.2 CPJ Notation

$$\mathbf{C} = \mathbf{H}^{-1} = \begin{bmatrix} c_{ii} & \mathbf{c}'_i \\ \mathbf{c}_i & \mathbf{C}_{[i]} \end{bmatrix}$$

? noted the following identity:

$$\mathbf{H}^{-1}_{[i]} = \mathbf{C}_{[i]} - \frac{1}{c_{ii}} \mathbf{c}_{[i]} \mathbf{c}'_{[i]}$$

? use the following as building blocks for case deletion statistics.

- \check{x}_i
- \check{z}_i
- $\check{z}_i j$
- \check{y}_i
- $p_i i$
- m_i

All of these terms are a function of a row (or column) of \mathbf{H} and $\mathbf{H}^{-1}_{[i]}$

0.6 CPJ's Three Propositions

Proposition 1

$$\mathbf{V}^{-1} = \begin{bmatrix} \nu^{ii} & \lambda'_i \\ \lambda_i & \Lambda_{[i]} \end{bmatrix}$$

$$\mathbf{V}_{[i]}^{-1} = \Lambda_{[i]} - \frac{\lambda_i \lambda'_i}{\lambda_i}$$

0.6.1 Proposition 2

$$(i) \quad \mathbf{X}_{[i]}^T \mathbf{V}_{[i]}^{-1} \mathbf{X}_{[i]} = \mathbf{X}' \mathbf{V}^{-1} \mathbf{X}$$

$$(ii) \quad = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{Y})^{-1}$$

$$(iii) \quad \mathbf{X}_{[i]}^T \mathbf{V}_{[i]}^{-1} \mathbf{Y}_{[i]} = \mathbf{X}' \mathbf{V}^{-1} \mathbf{Y}$$

0.6.2 Proposition 3

This proposition is similar to the formula for the one-step Newtown Raphson estimate of the logistic regression coefficients given by Pregibon (1981) and discussed in Cook Weisberg.

0.7 The CPJ Paper

0.7.1 Case-Deletion results for Variance components

? examines case deletion results for estimates of the variance components, proposing the use of one-step estimates of variance components for examining case influence. The method describes focuses on REML estimation, but can easily be adapted to ML or other methods.

This paper develops their global influences for the deletion of single observations in two steps: a one-step estimate for the REML (or ML) estimate of the variance components, and an ordinary case-deletion diagnostic for a weighted regression problem (conditional on the estimated covariance matrix) for fixed effects.

0.7.2 CPJ Notation

$$\mathbf{C} = \mathbf{H}^{-1} = \begin{bmatrix} c_{ii} & \mathbf{c}'_i \\ \mathbf{c}_i & \mathbf{C}_{[i]} \end{bmatrix}$$

? noted the following identity:

$$\mathbf{H}_{[i]}^{-1} = \mathbf{C}_{[i]} - \frac{1}{c_{ii}} \mathbf{c}_{[i]} \mathbf{c}'_{[i]}$$

? use the following as building blocks for case deletion statistics.

- \check{x}_i
- \check{z}_i
- \check{z}_{ij}
- \check{y}_i
- $p_i i$
- m_i

All of these terms are a function of a row (or column) of \mathbf{H} and $\mathbf{H}_{[i]}^{-1}$

0.8 Case Deletion Diagnostics

? develops case deletion diagnostics, in particular the equivalent of Cook's distance, for diagnosing influential observations when estimating the fixed effect parameters and variance components.

0.8.1 Deletion Diagnostics

Since the pioneering work of Cook in 1977, deletion measures have been applied to many statistical models for identifying influential observations.

Deletion diagnostics provide a means of assessing the influence of an observation (or groups of observations) on inference on the estimated parameters of LME models.

Data from single individuals, or a small group of subjects may influence non-linear mixed effects model selection. Diagnostics routinely applied in model building may identify such individuals, but these methods are not specifically designed for that purpose and are, therefore, not optimal. We describe two likelihood-based diagnostics for identifying individuals that can influence the choice between two competing models.

Case-deletion diagnostics provide a useful tool for identifying influential observations and outliers.

The computation of case deletion diagnostics in the classical model is made simple by the fact that estimates of β and σ^2 , which exclude the i th observation, can be computed without re-fitting the model. Such update formulas are available in the mixed model only if you assume that the covariance parameters are not affected by the removal of the observation in question. This is rarely a reasonable assumption.

0.9 Effects on fitted and predicted values

$$\hat{e}_{i(U)} = y_i - x\hat{\beta}_{(U)} \tag{1}$$

0.9.1 Case Deletion Diagnostics for Mixed Models

? notes the case deletion diagnostics techniques have not been applied to linear mixed effects models and seeks to develop methodologies in that respect.

? develops these techniques in the context of REML