0.0.1 Influence Statistics for LME models

Influence statistics can be coarsely grouped by the aspect of estimation that is their primary target:

- overall measures compare changes in objective functions: (restricted) likelihood distance (Cook and Weisberg 1982, Ch. 5.2)
- influence on parameter estimates: Cook's (Cook 1977, 1979), MDFFITS (Belsley, Kuh, and Welsch 1980, p. 32)
- influence on precision of estimates: CovRatio and CovTrace
- influence on fitted and predicted values: PRESS residual, PRESS statistic (Allen 1974), DFFITS (Belsley, Kuh, and Welsch 1980, p. 15)
- outlier properties: internally and externally studentized residuals, leverage

0.0.2 A Procedure for Quantifying Influence

The basic procedure for quantifying influence is simple as follows:

- Fit the model to the data and obtain estimates of all parameters.
- Remove one or more data points from the analysis and compute updated estimates of model parameters.
- Based on full- and reduced-data estimates, contrast quantities of interest to determine how the absence of the observations changes the analysis.

Cook (1986) introduces powerful tools for local-influence assessment and examining perturbations in the assumptions of a model. In particular the effect of local perturbations of parameters or observations are examined.

0.1 Measures 2

0.1.1 Cook's Distance

- For variance components γ

Diagnostic tool for variance components

$$C_{\theta i} = ((\hat{\theta})_{[i]} - (\hat{\theta}))^T \text{cov}((\hat{\theta}))^{-1}((\hat{\theta})_{[i]} - (\hat{\theta}))$$

0.1.2 Variance Ratio

- For fixed effect parameters β.

0.1.3 Cook-Weisberg statistic

- For fixed effect parameters β.

0.1.4 Andrews-Pregibon statistic

- For fixed effect parameters β.

The Andrews-Pregibon statistic AP_i is a measure of influence based on the volume of the confidence ellipsoid. The larger this statistic is for observation i, the stronger the influence that observation will have on the model fit.

0.2 Zewotir Measures of Influence in LME Models

Zewotir describes a number of approaches to model diagnostics, investigating each of the following;

- Variance components
- Fixed effects parameters

- Prediction of the response variable and of random effects
- likelihood function

Random Effects

A large value for $CD(u)_i$ indicates that the $i-$th observation is influential in predicting random effects.

linear functions

$CD(\psi)_i$ does not have to be calculated unless $CD(\beta)_i$ is large.

0.2.1 Information Ratio

0.3 Computation and Notation

with $\boldsymbol{V}$ unknown, a standard practice for estimating $\boldsymbol{X\beta}$ is the estime the variance components σ_j^2, compute an estimate for $\boldsymbol{V}$ and then compute the projector matrix A, $\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{AY}$.

Zewotir remarks that $\boldsymbol{D}$ is a block diagonal with the $i-$th block being $u\boldsymbol{I}$