Haslett and Hayes - Residuals

Haslett and Hayes (1998) and Haslett (1999) considered the case of an LME model with correlated covariance structure.

0.0.1 Residual Diagnostics in LME models

- A **residual** is the difference between the observed quantity and the predicted value. In LME models a distinction is made between marginal residuals and conditional residuals.

- A **Marginal residual** is the difference between the observed data and the estimated marginal mean (Schabenberger pg3) The computation of case deletion diagnostics in the classical model is made simple by the fact that important estimates can be computed without refitting the model.

- Such update formulae are available in the mixed model only if you assume that the covariance parameters are not affect by the removal of the observation in question. Schabenberger remarks that this is not a reasonable assumption.

Basic procedure for quantifying influence is simple

1. Fit the model to the data

2. Remove one or more data points from the analysis and compute updated estimates of model parameters

3. Based on the full and reduced data estimates, contrast quantities of interest to determine how the absence of the observations changed the analysis.

The likelihood distance is a global summary measure expressing the joint influence of the observations in the set U on all parameters in Ψ that were subject to updating.

0.1 Case Deletion Diagnostics for LME models

? remark that linear mixed effects models didn't experience a corresponding growth in the use of deletion diagnostics, adding that **?** makes no mention of diagnostics whatsoever.

? describes three propositions that are required for efficient case-deletion in LME models. The first proposition decribes how to efficiently update V when the ith element is deleted.

$$V_{[i]}^{-1} = \Lambda_{[i]} - \frac{\lambda\lambda\prime}{\nu ii} \tag{1}$$

The second of christensen's propostions is the following set of equations, which are variants of the Sherman Wood bury updating formula.

$$X'_{[i]}V_{[i]}^{-1}X_{[i]} = X'V^{-1}X - \frac{\hat{x}_i\hat{x}'_i}{s_i} \tag{2}$$

$$(X'_{[i]}V_{[i]}^{-1}X_{[i]})^{-1} = (X'V^{-1}X)^{-1} + \frac{(X'V^{-1}X)^{-1}\hat{x}_i\hat{x}'_i(X'V^{-1}X)^{-1}}{s_i - \bar{h}_i} \tag{3}$$

$$X'_{[i]}V_{[i]}^{-1}Y_{[i]} = X\prime V^{-1}Y - \frac{\hat{x}_i\hat{y}'_i}{s_i} \tag{4}$$

In LME models, fitted by either ML or REML, an important overall influence measure is the likelihood distance (**?**). The procedure requires the calculation of the full data estimates $\hat{\psi}$ and estimates based on the reduced data set $\hat{\psi}_{(U)}$. The likelihood distance is given by determining

$$LD_{(U)} = 2\{l(\hat{\psi}) - l(\hat{\psi}_{(U)})\} \tag{5}$$

$$RLD_{(U)} = 2\{l_R(\hat{\psi}) - l_R(\hat{\psi}_{(U)})\} \tag{6}$$