

Chapter 1

Residuals for LME Models

1.1 Residual diagnostics

In classical linear models model diagnostics have been become a required part of any statistical analysis, and the methods are commonly available in statistical packages and standard textbooks on applied regression. A visual inspection for the presence of trends inform the analyst on the validity of distributional assumptions, and to detect outliers and influential observations.

However it has been noted by several papers that model diagnostics do not often accompany LME model analyses.

1.1.1 LME REsiduals

Cox and Snell (1968, JRSS-B): general definition of residuals for models with single source of variability Hilden-Minton (1995, PhD thesis UCLA), Verbeke and Lesaffre (1997, CSDA) or Pinheiro and Bates (2000, Springer): extension to define three types of residuals that accommodate the extra source of variability present in linear mixed models, namely:

1. Marginal residuals,

predictors of marginal errors,

2. Conditional residuals,

$$be = yX\hat{\beta}Zbb = \hat{\sigma}Q\hat{y}$$

, predictors of conditional errors

$$e = yE[y|b] = yX\beta Zb$$

3. BLUP, Zbb , predictors of random effects,

$$Zb = E[y|b]E[y]$$

1.1.2 Residual Diagnostics in LME models

- A **residual** is the difference between the observed quantity and the predicted value.
- In LME models you can distinguish marginal residuals rm and conditional residuals rc . A marginal residual is the difference between the observed data and the estimated (marginal) mean.
- 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.
- Residuals are used to examine model assumptions and to detect outliers and potentially influential data point. The raw residuals r_{mi} and r_{ci} are usually not well suited for these purposes.

- In LME models, there are two types of residuals, marginal residuals and conditional residuals. In a model without random effects, both sets of residuals coincide. schabenberger provides a useful summary.
- 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.
- A marginal residual is the difference between the observed data and the estimated (marginal) mean, $r_{mi} = y_i - x'_0 \hat{b}$
- A conditional residual is the difference between the observed data and the predicted value of the observation, $r_{ci} = y_i - x'_i \hat{b} - z'_i \hat{\gamma}$
- 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.
- The marginal and conditional means in the linear mixed model are $E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$ and $E[\mathbf{Y}|\mathbf{u}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$, respectively.
- In linear mixed effects models, diagnostic techniques may consider ‘conditional’ residuals. A conditional residual is the difference between an observed value y_i and the conditional predicted value \hat{y}_i .

$$\hat{\epsilon}_i = y_i - \hat{y}_i = y_i - (X_i \hat{\beta} + Z_i \hat{b}_i)$$

However, using conditional residuals for diagnostics presents difficulties, as they tend to be correlated and their variances may be different for different subgroups,

which can lead to erroneous conclusions.

$$r_{mi} = x_i^T \hat{\beta} \quad (1.1)$$

- Conditional residuals include contributions from both fixed and random effects, whereas marginal residuals include contribution from only fixed effects.

Marginal residuals

$$y - X\beta = Z\eta + \epsilon$$

- Should be mean 0, but may show grouping structure
- May not be homoskedastic.
- Good for checking fixed effects, just like linear regr.

1.1.3 Marginal Residuals

$$\begin{aligned} \hat{\beta} &= (X^T R^{-1} X)^{-1} X^T R^{-1} Y \\ &= BY \end{aligned}$$

Marginal Residuals

Distinction From Linear Models

- The differences between perturbation and residual analysis in the linear model and the linear mixed model are connected to the important facts that b and b depend on the estimates of the covariance parameters, that b has the form of an (estimated) generalized least squares (GLS) estimator, and that is a random vector.

- In a mixed model, you can consider the data in a conditional and an unconditional sense. If you imagine a particular realization of the random effects, then you are considering the conditional distribution Y —
- If you are interested in quantities averaged over all possible values of the random effects, then you are interested in Y ; this is called the marginal formulation. In a clinical trial, for example, you may be interested in drug efficacy for a particular patient. If random effects vary by patient, that is a conditional problem. If you are interested in the drug efficacy in the population of all patients, you are using a marginal formulation. Correspondingly, there will be conditional and marginal residuals, for example.
- The estimates of the fixed effects

depend on the estimates of the covariance parameters. If you are interested in determining the influence of an observation on the analysis, you must determine whether this is influence on the fixed effects for a given value of the covariance parameters, influence on the covariance parameters, or influence on both.
- Mixed models are often used to analyze repeated measures and longitudinal data. The natural experimental or sampling unit in those studies is the entity that is repeatedly observed, rather than each individual repeated observation. For example, you may be analyzing monthly purchase records by customer.
- An influential data point is then not necessarily a single purchase. You are probably more interested in determining the influential customer. This requires that you can measure the influence of sets of observations on the analysis, not just influence of individual observations.
- The computation of case deletion diagnostics in the classical model is made simple

by the fact that 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.

- The application of well-known concepts in model-data diagnostics to the mixed model can produce results that are at first counter-intuitive, since our understanding is steeped in the ordinary least squares (OLS) framework. As a consequence, we need to revisit these important concepts, ask whether they are portable to the mixed model, and gain new appreciation for their changed properties. An important example is the ostensibly simple concept of leverage.
- The definition of leverage adopted by the MIXED procedure can, in some instances, produce negative values, which are mathematically impossible in OLS. Other measures that have been proposed may be non-negative, but trade other advantages. Another example are properties of residuals. While OLS residuals necessarily sum to zero in any model (with intercept), this not true of the residuals in many mixed models.

1.1.4 Confounded Residuals

Hilden-Minton (1995, PhD thesis, UCLA): residual is pure for a specific type of error if it depends only on the fixed components and on the error that it is supposed to predict. Residuals that depend on other types of errors are called *confounded residuals*.

1.2 Conditional and Marginal Residuals

Suppose the linear mixed-effects model lme has an $n \times p$ fixed-effects design matrix \mathbf{X} and an $n \times q$ random-effects design matrix \mathbf{Z} .

Also, suppose the p-by-1 estimated fixed-effects vector is $\hat{\beta}$, and the q-by-1 estimated best linear unbiased predictor (BLUP) vector of random effects is \hat{b} . The fitted conditional response is

$$\hat{y}_{Cond} = X\hat{\beta} + Z\hat{b}$$

and the fitted marginal response is

$$\hat{y}_{Mar} = X\hat{\beta}$$

residuals can return three types of residuals:

- raw,
- Pearson, and
- standardized.

For any type, you can compute the conditional or the marginal residuals. For example, the conditional raw residual is

$$r_{Cond} = y - X\hat{\beta} - Z\hat{b}$$

and the marginal raw residual is

$$r_{Mar} = y - X\hat{\beta}$$

Conditional residuals

$$y - X\beta - Z\eta = \epsilon$$

- Should be mean zero with no grouping structure

- Should be homoscedastic.
- Good for checking normality of outliers

Random effects

$$y - X\beta - \epsilon = Z\eta$$

- Should be mean zero with no grouping structure
- May not be homoscedastic.

1.3 Computation and Notation

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

Zewotir and Galpin (2005) remarks that \mathbf{D} is a block diagonal with the i -th block being $u\mathbf{I}$

1.4 Residual Diagnostics

Consider a residual vector of the form $\hat{e} = \mathbf{P}\mathbf{Y}$, where \mathbf{P} is a projection matrix, possibly an oblique projector. External studentization uses an estimate of Var that does not involve the i th observation.

Externally studentized residuals are often preferred over studentized residuals because they have well known distributional properties in the standard linear models for independent data.

Residuals that are scaled by the estimated variances of the responses are referred to as Pearson-type residuals.

Standardization:

$$\frac{\hat{e}_i}{\sqrt{v_i}}$$

Studentization

$$\frac{\hat{e}_i}{\sqrt{\hat{v}_i}}$$

1.5 Marginal Residuals

$$\begin{aligned}\hat{\beta} &= (X^T R^{-1} X)^{-1} X^T R^{-1} Y \\ &= BY\end{aligned}$$

1.5.1 Computation

The computation of internally studentized residuals relies on the diagonal entries of $V(\hat{\theta}) - Q(\hat{\theta})$, where $Q(\hat{\theta})$ is computed as

$$Q(\hat{\theta}) = X(X'Q(\hat{\theta})^{-1}X)X^{-1}$$

1.5.2 Pearson Residual

Another possible scaled residual is the ‘Pearson residual’, whereby a residual is divided by the standard deviation of the dependent variable. The Pearson residual can be used when the variability of $\hat{\beta}$ is disregarded in the underlying assumptions.

1.6 Effects on fitted and predicted values

$$\hat{e}_{i(U)} = y_i - x\hat{\beta}_{(U)} \tag{1.2}$$

1.7 Diagnostic Tools for the nlme package

With the nlme package, the generic function `lme()` fits a linear mixed-effects model in the formulation described in Laird and Ware (1982) but allowing for nested random effects.

The within-group errors are allowed to be correlated and/or have unequal variances, which is very important in fitting the models for Roy's Tests

The nlme package has a limited set of diagnostic tools that can be used to assess the model fit. A review of the package manual is sufficient to get a sense of the package's capability in that regard.

1.7.1 `residuals.lme` nlme- Extract lme Residuals

The residuals at level i are obtained by subtracting the fitted levels at that level from the response vector (and dividing by the estimated within-group standard error, if `type="pearson"`).

The fitted values at level i are obtained by adding together the population fitted values (based only on the fixed effects estimates) and the estimated contributions of the random effects to the fitted values at grouping levels less or equal to i .

```
fm1 <- lme(distance ~ age + Sex,  
data = Orthodont, random = ~ 1)  
head(residuals(fm1, level = 0:1))  
summary(residuals(fm1) /  
residuals(fm1, type = "p"))
```

```
# constant scaling factor 1.432
```

Bibliography

Cook, R. (1986). Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)* 48(2), 133–169.

Zewotir, T. and J. S. Galpin (2005). Influence diagnostics for linear mixed models. *Journal of Data Science* 3(2), 153–177.