Supplementary material for "Negative variance components and intercept-slope correlations greater than one in magnitude: How do such 'non-regular' random intercept and slope models arise, and what should be done when they do?" by Helen Bridge, Katy E Morgan, and Chris Frost

Technical Appendix

1 Eigenvalues and eigenvectors of $\boldsymbol{ZGZ}^T$

From eigenvalue theory it is known that if $\boldsymbol{A}$ is an *r*-by-*s* matrix and $\boldsymbol{B}$ is an *s*-by-*r* matrix, such that $s \geq r$, then the *s* eigenvalues of $\boldsymbol{BA}$ are the *r* eigenvalues of $\boldsymbol{AB}$, with the additional $s - r$ eigenvalues being zero. It follows that $\boldsymbol{ZGZ}^T$ has only two non-zero eigenvalues, these being the eigenvalues of the 2-by-2 matrix $\boldsymbol{GZ^TZ}$, with the remaining $n - 2$ eigenvalues all being zero.

Because $\boldsymbol{Z^T v}_i = \boldsymbol{0}$ guarantees that $\boldsymbol{ZGZ^T v}_i = \boldsymbol{0}$, the $n - 2$ non-defining eigenvectors (the eigenvectors with eigenvalues equal to zero) will be an arbitrary set of vectors orthogonal to the column vectors that make up $\boldsymbol{Z}$ (a column of 1's and a column of measurement times t_1 to t_n), while the eigenvectors that correspond to the defining eigenvalues (which we term the defining eigenvectors) will both be linear combinations of the column vectors that make up $\boldsymbol{Z}$.

2 'Defining eigenvalues' and standard errors of $\widehat{\boldsymbol{\beta}}$

The defining eigenvalues of $\boldsymbol{\Sigma}$ also have relevance for the standard errors of $\widehat{\boldsymbol{\beta}}$ obtained when fitting the RIAS model to data. From standard theory $\boldsymbol{\Sigma}$ can be written as $\boldsymbol{V\Lambda V}^T$ where $\boldsymbol{V}$ denotes the matrix of eigenvectors of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}$. Since $\boldsymbol{V}$ is a matrix of eigenvectors, $\boldsymbol{V}^{-1} = \boldsymbol{V}^T$ and so

$$\boldsymbol{\Sigma}^{-1} = (\boldsymbol{V\Lambda V}^T)^{-1} = \boldsymbol{V\Lambda}^{-1}\boldsymbol{V}^T. \qquad \text{(T1)}$$

For a mixed model written (as in equation (2)) as $\boldsymbol{Y_i}|\boldsymbol{b_i} \sim N(\boldsymbol{X\beta} + \boldsymbol{Zb_i},\ \boldsymbol{R_2})$,

$$V(\widehat{\boldsymbol{\beta}}) = \frac{1}{N}(\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}. \qquad \text{(T2)}$$

Hence,

$$V(\widehat{\boldsymbol{\beta}}) = \frac{1}{N}(\boldsymbol{X}^T\boldsymbol{V\Lambda}^{-1}\boldsymbol{V}^T\boldsymbol{X})^{-1}. \qquad \text{(T3)}$$

For the RIAS model in equation (1) $\boldsymbol{X} = \boldsymbol{Z}$ and so, because (as shown in Section 1) the non-defining eigenvectors of $\boldsymbol{\Sigma}$ are orthogonal to $\boldsymbol{X}$, the matrix $\boldsymbol{X}^T\boldsymbol{V}$ has all but two columns that are made up of zeros. Analogously $\boldsymbol{V}^T\boldsymbol{X}$ has all but two rows that are made up of zeros. Further, these zero columns and rows multiply the reciprocal of the repeated, non-defining eigenvalue in $\boldsymbol{\Lambda}^{-1}$. Hence the standard errors of $\widehat{\boldsymbol{\beta}}$ depend on the defining eigenvalues of $\boldsymbol{\Sigma}$ (θ_1 and θ_2) but not on the non-defining eigenvalues.

3 Impact of changing the number and timing of follow-up visits

The form of the PSD defining matrix $\boldsymbol{D}$ has implications concerning the number and timing of measurements. These follow from a theory due to Ostrowski[1]. The theory relates to the situation where $\boldsymbol{A}$ is a symmetric 2-by-2 matrix with eigenvalues λ_1 and λ_2 and $\boldsymbol{B}$ is a 2-by-2 diagonal matrix with positive elements ($c_{max} \geq c_{min} > 0$). Ostrowski proves that the eigenvalues of $\boldsymbol{AB}$ are $d_1\lambda_1$ and $d_2\lambda_2$ where $c_{max} \geq d_i \geq c_{min}$ for $i = 1, 2$.

Now consider the RIAS model and suppose that we decrease n to n^* and q to q^* while keeping $\boldsymbol{G}$ constant. The effect of this is to post multiply $\boldsymbol{GZ}^T\boldsymbol{Z}$ by a 2-by-2 diagonal matrix both of whose elements are less than 1 (n^*/n and q^*/q respectively). By the Ostrowski theory each of the defining eigenvalues of $\boldsymbol{ZGZ}^T$ (the eigenvalues of $\boldsymbol{GZ}^T\boldsymbol{Z}$) is multiplied by a number that lies between n^*/n and q^*/q, so reducing their magnitude. If the defining eigenvalues are initially positive, or negative but not less than $-\sigma_e^2$, they will remain not less than $-\sigma_e^2$. However, if we increase n and q while keeping $\boldsymbol{G}$ constant, we cannot guarantee that the defining eigenvalues will stay not less than $-\sigma_e^2$. If the defining eigenvalues are both initially non-negative then they will stay non-negative, but if one or both of them is negative then adding follow-up visits will increase the magnitude of that eigenvalue, potentially such that it is greater than σ_e^2 in magnitude. Indeed, when $\boldsymbol{ZGZ}^T$ has at least one negative eigenvalue, it is not possible to continually increase n and q while $\boldsymbol{G}$ remains unchanged: ultimately this must result in $\boldsymbol{\Sigma}$ no longer remaining PSD.

4 Non-linearity in fixed effects as a cause of non-regularity

Imagine that we observe y_{ij}^*, where $y_{ij}^* = y_{ij} + f_j$ and y_{ij} follows a simple RIAS model (as defined in equation (1)) that is regular, and $f_j = \sum_{k=2}^{n-1} \beta_k f_k(t_j)$ where each $f_k(t_j)$ is a polynomial function of time discretely orthogonal both to a constant and to linear time. For example, with data at five time points ($t_1 = -2,\ t_2 = -1,\ t_3 = 0,\ t_4 = 1,\ t_5 = 2$) one such set of discrete orthogonal polynomials is $f_2(t_j) = t_j^2 - 2,\ f_3(t_j) = (5t_j^3 - 17t_j)/6$ and $f_4(t_j) = (35t_j^4 - 155t_j^2 + 72)/12$, these giving the vectors (2, -1,-2, -1, 2)^T, (-1, 2, 0, -2, 1)^T and (1, -4, 6, -4, 1)^T respectively, which are mutually orthogonal and also orthogonal to (-2, -1, 0, 1, 2)^T (ie, a vector of linear time) and (1, 1, 1, 1, 1)^T (ie, a constant). Adding these three polynomials to the constant and linear terms puts no constraint on the means at the five time points, so any non-linearity in the relationship between the mean of the outcome and time can be accommodated.

Now contrast fitting the RIAS model in equation (1) to the observed y_{ij}^* rather than to y_{ij}. The y_{ij}^* do not follow the RIAS model, but because of the orthogonality of the $f_k(t_j)$ functions, the estimates of the fixed linear and constant terms, β_0 and β_1, will be the same whether y_{ij}^* or y_{ij} is modelled. However, modelling y_{ij}^* rather than y_{ij} with the RIAS model will cause the expectation of the estimate of $\boldsymbol{\Sigma} = \boldsymbol{R_n} + \boldsymbol{ZGZ}^T$ to change because there is now additional variability not accounted for by the fixed effects in the model. Further, the effect on $\boldsymbol{\Sigma} = \boldsymbol{R_n} + \boldsymbol{ZGZ}^T$ is predictable because this additional variability is orthogonal to $\boldsymbol{Z}$. Specifically, the defining eigenvalues of $\boldsymbol{\Sigma}$ will remain unchanged whilst the non-defining eigenvalue (ie, the residual variance represented by the diagonal elements of $\boldsymbol{R}$) will increase in expectation by an amount equal to the residual variance from a simple linear regression of f_j on t_j.

If this increase in the non-defining eigenvalue of $\boldsymbol{\Sigma}$ is such that it remains smaller than the other two (defining) eigenvalues, then fitting the RIAS model to y_{ij}^* will give parameter estimates that correspond to a regular RIAS model. However, if the non-defining eigenvalue becomes larger than either of the other two, then parameter estimates that correspond to a non-regular RIAS model can result.

5 RIAS and random quadratic models for data at three evenly spaced time-points

For data at three evenly spaced time points, a number of mixed models that include all the terms in the simple RIAS model in equation (1) plus an additional random quadratic term all have the same marginal variance-covariance matrix. Specifically, all models parameterized as

$$y_{ij} = \beta_0 + \beta_1 t_j + b_{0i} + b_{1i} t_j + b_{2i} t_j^2 + e_{ij} : \ t_1 = -1, t_2 = 0, t_3 = 1$$

where $\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b0}^2 + 2k & \sigma_{b01} & -2k \\ \sigma_{b01} & \sigma_{b1}^2 + k & 0 \\ -2k & 0 & 3k \end{pmatrix}\right]; e_{ij} \sim N[0, \sigma_e^2 - 2k]$

have $\mathrm{Var}\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} = \begin{pmatrix} \sigma_{b0}^2 + \sigma_{b1}^2 - 2\sigma_{b01} + \sigma_e^2 & \sigma_{b0}^2 - \sigma_{b01} & \sigma_{b0}^2 - \sigma_{b1}^2 \\ \sigma_{b0}^2 - \sigma_{b01} & \sigma_{b0}^2 + \sigma_e^2 & \sigma_{b0}^2 + \sigma_{b01} \\ \sigma_{b0}^2 - \sigma_{b1}^2 & \sigma_{b0}^2 + \sigma_{b01} & \sigma_{b0}^2 + \sigma_{b1}^2 + 2\sigma_{b01} + \sigma_e^2 \end{pmatrix}$

for all choices of $k \leq \sigma_e^2/2$.

This demonstrates that if three-point data are compatible with the RIAS model ($k = 0$) then they are also compatible with a whole set of parameterizations of the 'random quadratic model'. So, we can think of non-regular RIAS model three-point data as being generated by a random quadratic model with parameters that cannot be uniquely estimated from the data. Further, one such parameterization is $k = \sigma_e^2/2$, which implies that $\boldsymbol{\Sigma} = \boldsymbol{ZGZ}^T$ and so (by the Ostrowski rule referred to in Section 3 above) if $\boldsymbol{\Sigma}$ is PSD then $\boldsymbol{G}$ will be too. So, a non-regular RIAS model for three-point data has at least one random quadratic model analogue that is regular.

6 Code for models fitted to the rat data[2, 3]

6.1 SAS Code

```
/* Transform age to create time variable as modelled by
Molenberghs and Verbeke (2000)*/
data rats;
  set rats.rats;
  time = log(1 + (age - 45) / 10);
run;

title "Random intercept and slope model, with 'nobound'
option";
proc mixed data = rats covtest nobound;
  class treat;
  model response = treat * time / solution;
  repeated / type = simple subject = rat r;
  random intercept time / type = un subject = rat g gcorr;
run;
title;

title "Random intercept and slope model model, without
'nobound'";
proc mixed data = rats covtest;
```

```
  class treat;
  model response = treat * time / solution;
  repeated / type = simple subject = rat r;
  random intercept time / type = un subject = rat g gcorr;
run;
title;

title "Random intercept model";
proc mixed data = rats covtest;
  class treat;
  model response = treat * time / solution;
  repeated / type = simple subject = rat r;
  random intercept / type = un subject = rat g gcorr;
run;
title;
```

6.2 Stata code

```
use "rats", clear

* Transform age to create time variable as modelled by
Molenberghs and Verbeke (2000)
gen time = log(1 + (age – 45) / 10)

* Convert treatment variable from string to numeric form
encode treat, gen(trt)

* Fit RIAS model with default output (random effects variances
and covariance)
mixed response i.tr#c.time || rat: time, reml
cov(unstructured) residuals(independent)

* Fit RIAS model with 'stddev' option for output including
random effects standard deviations and correlation
mixed response i.tr#c.time || rat: time, ///
     reml cov(unstructured) residuals(independent) stddev
```

6.3 R code

```
library(lme4)
rats <- read.csv("rats.csv")

# Transform age to create time variable as modelled by
Molenberghs and Verbeke (2000)
rats$time <- log(1 + (rats$age – 45) / 10)

# Fit RIAS model
```

```
model <- lmer(response ~ treat: time + (1 + time | rat), data
= rats)
summary(model)

# Confirm that model is classed as a boundary (singular) fit
isSingular(model)
```

7 References

1. Horn RAJ, Charles R. Matrix Analysis. 2nd ed. Cambridge: Cambridge University Press; 2012.

2. Molenberghs G, Verbeke G. Linear Mixed Models for Longitudinal Data. New York: Springer; 2000.

3. Verdonck A, De Ridder L, Kühn R, Darras V, Carels C, de Zegher F. Effect of testosterone replacement after neonatal castration on craniofacial growth in rats. Arch Oral Biol. 1998;43(7):551-557. Dataset available at: https://gbiomed.kuleuven.be/english/research/50000687/50000696/geertverbeke/datasets. Accessed November 5, 2023.