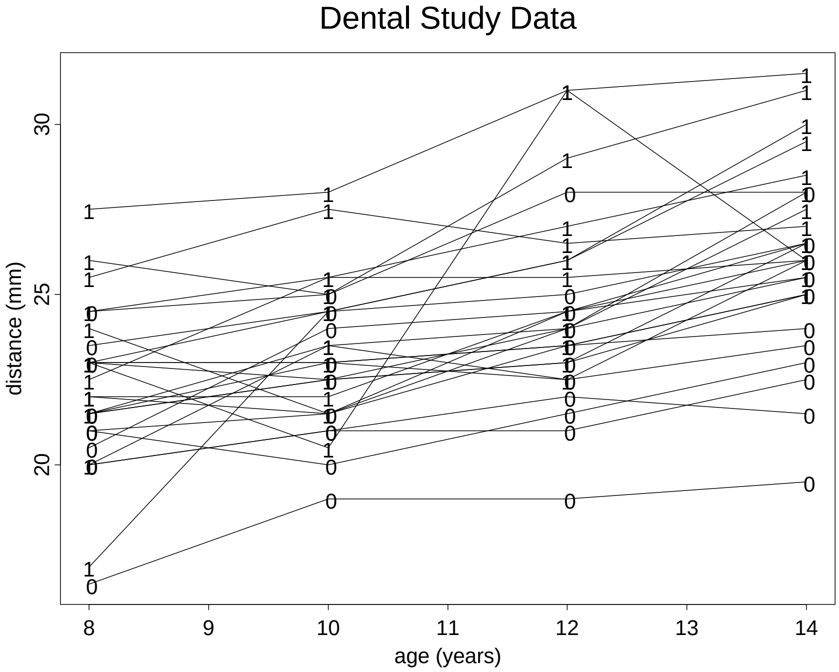
9 Random coefficient models for multivariate normal data

9.1 Introduction

In the last chapter, we noted that an alternative perspective on explicit modeling of longitudinal response is to think directly of the fact that each unit appears to have its **own trajectory** or **inherent trend** with its own peculiar features. For example, in the dental study, if we focus on a particular child, the trajectory looks to be approximately like a straight line (with some variation about it, of course). The data are reproduced below for convenience in Figure 1. A similar statement could be made about the dialyzer data in the last chapter.

Figure 1: *Dental data revisited.*



The general regression modeling approach takes the standard perspective in much of statistical modeling of focusing directly on the **mean responses** and how they change over time. In this chapter, we consider an alternative approach to building a model based on thinking first about individual trajectories.

- For trajectories that may be represented by **linear functions** of a design matrix and parameters, this approach will lead us to the same type of mean models as the general regression approach.

- **However**, the modeling approach acknowledges **explicitly** the two separate sources of variation we have discussed. As a result, it "automatically" leads to covariance models that also acknowledge these sources.

- The resulting statistical model, called a **random coefficient model** for reasons that will be clear shortly, will be seen to imply a a model like the general linear regression models of the last chapter with a particular covariance structure for each data vector. Thus, the inferential methods of that chapter, namely maximum and restricted maximum likelihood, will apply immediately.

- In addition, this modeling strategy will allow us to address questions of scientific interest about trajectories for **individual units**, either ones in the study or **future** units. For example, in a study of AIDS patients, it may be of interest to physicians attending the patients to have an **estimate** of a patient's individual apparent trajectory, so that they may make clinical decisions about his or her future care. There is no apparent way of doing this in the general modeling approach we have just considered.

9.2 Random coefficient model

SUBJECT-SPECIFIC TRAJECTORY: Recall the conceptual model discussed in Chapter 4. For definiteness, again consider the dental study data. We take the view that each child has his/her own underlying straight line **inherent trend**. Focusing on the ith child, this says that s/he has his/her own **intercept** and **slope**, β_{0i} and β_{1i}, say, respectively, that determine this trend. This intercept and slope are unique to child i.

WITHIN-INDIVIDUAL VARIATION: Continuing with conceptual perspective, the actual responses observed for a given child do not fall **exactly** on a straight line (the inherent trajectory) due to

- The fact that the response cannot be measured perfectly, but is instead subject to measurement error due to the measuring device.

- Individual "fluctuations;" although the overall **trend** for a given child is a straight line, the **actual responses**, if we could observe them continuously over time, tend to fluctuate about the trend.

AMONG-INDIVIDUAL VARIATION: The inherent trajectories are "high" or "low" with different steepness across children, suggesting that the child-specific **intercepts** β_{0i} and **slopes** β_{1i} **vary** across children.

To formalize this thinking, a model is developed in two **stages**.

"INDIVIDUAL (FIRST STAGE)" MODEL: The first stage involves describing what we believe at the level the ith child; specifically, we write a model for the random variables $Y_{i1}, \ldots, Y_{in_i}$ for the ith child taken at time points $t_{i1}, \ldots, t_{in_i}$. Although the particular dental study example is **balanced**, we write things more generally to allow the possibility of imbalance. The model for child i is, $i = 1, \ldots, m$ is

$$Y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} + e_{ij}, \quad j = 1, \ldots, n_i. \tag{9.1}$$

In model (9.1), the observations on the ith child follow a straight line with child-specific intercept and slope β_{0i} and β_{1i}. That actual observations vary about this inherent line due to within-unit sources is represented explicitly by the deviation e_{ij} with mean 0. We say more about these deviations shortly.

- Thus, model (9.1) has the form of a straight line regression model **unique** to the ith child. Each child has such a model.
- Each child has a **regression parameter** vector $\boldsymbol{\beta}_i = \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \end{pmatrix}$.
- We may write the model (9.1) concisely. Define $\boldsymbol{Y}_i$ and $\boldsymbol{e}_i$ as usual, and let

$$\boldsymbol{Z}_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \\ 1 & t_{in_i} \end{pmatrix}.$$

We may then write the model as

$$\boldsymbol{Y}_i = \boldsymbol{Z}_i \boldsymbol{\beta}_i + \boldsymbol{e}_i, \quad i = 1, \ldots, m. \tag{9.2}$$

"POPULATION (SECOND STAGE)" MODEL: Model (9.1) only tells part of the story; it describes what happens at the level of an individual child, and includes explicit mention (through e_{ij}) of **within-child** variation. However, it does not by itself acknowledge **among-child** variation. We have recognized that the inherent trends differ across children; for example, some children have a steeper slope for their apparent trajectory than do others. For now, we downplay the fact that children are of two genders; we will tackle this issue momentarily.

We may think of the children observed as arising from a **population** of all such children. Each child has its **own** intercept and slope; thus, we may think abstractly of this population in terms of **random vectors** $\boldsymbol{\beta}_i$, one for each child, as it is the unique intercept and slope for each child that distinguishes his/her trajectory.

- It is natural to think of this **population** as being "centered" about a "typical" value of intercept and slope, with variation about this center value – some children have shallower or steeper slopes, for example.

- More formally, we may think of the **mean** value of intercept and slope of the population of all such $\boldsymbol{\beta}_i$ vectors. Individual intercept/slope vectors vary about this mean. Thus, we may think of a **joint probability distribution** of all possible values that a random vector of regression parameters $\boldsymbol{\beta}_i$ could take on. More on this momentarily.

This way of thinking suggests a **model** for this population as follows. Let β_0 and β_1 represent the **mean** values of intercept and slope, and define

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}. \tag{9.3}$$

Thus $\boldsymbol{\beta}$ is the **mean vector** of the population of all $\boldsymbol{\beta}_i$. Then write

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{b}_i, \quad \boldsymbol{b}_i = \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix}, \tag{9.4}$$

which is a shorthand way of saying

$$\beta_{0i} = \beta_0 + b_{0i}, \quad \beta_{1i} = \beta_1 + b_{1i}.$$

- Here, $\boldsymbol{b}_i$ is a vector of **random effects** describing how the intercept and slope for the ith child deviates from the mean value.

- Thus, (9.4) has the flavor of a regression-type model for the child-specific regression parameters, with a **systematic** component, the **mean**, and a **random** component summarizing how things vary about it.

- More formally, the vectors $\boldsymbol{b}_i$ are assumed to have mean $\mathbf{0}$ and some **covariance matrix** that describes the nature of this variation – how intercepts and slopes vary **among** children **and** how they **covary** (e.g. do large intercepts and slopes tend to occur together?) In fact, as we discuss shortly, the $\boldsymbol{b}_i$ are assumed to have a **multivariate probability distribution** with this mean and covariance matrix.

- Thus, whereas the **individual** child model summarizes how things happen **within** a child, this model characterizes variation **among** children, representing the population through intercepts and slopes. Putting the models (9.1) and (9.4) **together** thus gives a complete description of what we believe about each child and the population of children, acknowledging the two sources of variation **explicitly**.

- Note that we may substitute the expressions for β_{0i} and β_{1i} in (9.1) to obtain

  $$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + e_{ij}.$$

  This shows clearly what we are assuming: each child has intercept and slope that varies about the "typical," or **mean** intercept and slope β_0 and β_1.

ACKNOWLEDGING GENDER: We can refine our model to allow for the fact that children are of different genders as follows. We may think of children as coming from two **populations**, males and females, each population with its own **mean** values of intercept and slope and possibly **different** pattern of variation in these intercepts and slopes. Each child would still have his/her own individual regression model as in (9.1), so this would not change. What would change to incorporate this refinement is the **population model**. For example, if child i is a boy, then we might believe

$$\beta_{0i} = \beta_{0,B} + b_{0i}. \;\; \beta_{1i} = \beta_{1,B} + b_{1i},$$

while if i is a girl,

$$\beta_{0i} = \beta_{0,G} + b_{0i}. \;\; \beta_{1i} = \beta_{1,G} + b_{1i}.$$

- Here, the **fixed** parameters $\beta_{0,B}, \beta_{1,B}$ represent the mean intercept and slope for boys; similarly, $\beta_{0,G}, \beta_{1,G}$ represent the same for girls.

- $\boldsymbol{b}_i = (b_{0i}, b_{1i})'$ represents the **random effect** for child i with mean **0** We may believe that the populations of $\boldsymbol{\beta}_i$ for boys and girls have different means but have similar variation. In this case, we might say that the $\boldsymbol{b}_i$ all have the **same** covariance matrix regardless of whether i is a boy or girl. On the other hand, if we believe that the populations have different variation, we might think of the $\boldsymbol{b}_i$ of being of two types, with a different covariance matrix depending on the gender. We will be more formal shortly.

- Let

  $$\boldsymbol{\beta} = \begin{pmatrix} \beta_{0,G} \\ \beta_{1,G} \\ \beta_{0,B} \\ \beta_{1,B} \end{pmatrix}.$$

Define for each child a matrix $\boldsymbol{A}_i$ such that

$$\boldsymbol{A}_i = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ if child } i \text{ is a girl}$$

$$\boldsymbol{A}_i = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ if child } i \text{ is a boy}$$

Then it is straightforward to verify that we may write the model concisely for each child as

$$\boldsymbol{\beta}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i. \tag{9.5}$$

- Note that the simpler ("one-population") model (9.4) could also be written in this way with $\boldsymbol{\beta}$ defined as in (9.3) and $\boldsymbol{A}_i = \boldsymbol{I}_2$ for all i (try it!)

Let us now be more specific about the nature of the two sources of variation being acknowledged explicitly in this modeling approach.

WITHIN-UNIT VARIATION: In the "individual" model (9.2), the **within-unit** random vector $\boldsymbol{e}_i$ has mean zero and represents the deviations introduced **solely by** sources within an individual. This includes measurement error, biological "fluctuations," or both. Thus, following the conceptual framework in Chapter 4, we may think of $\boldsymbol{e}_i$ as being decomposed as

$$\boldsymbol{e}_i = \boldsymbol{e}_{1i} + \boldsymbol{e}_{2i},$$

where $\boldsymbol{e}_{1i}$ represents the deviations due to within-subject fluctuations and $\boldsymbol{e}_{2i}$ those due to measurement error.

To characterize within-subject variation **and** correlation due to within-subject sources (fluctuations), the approach is to specify a **covariance structure model** for $\text{var}(\boldsymbol{e}_i)$. In general, write

$$\boldsymbol{R}_i = \text{var}(\boldsymbol{e}_i),$$

where $\boldsymbol{R}_i$ is a $(n_i \times n_i)$ covariance matrix. We now discuss through review of some typical scenarios considerations involved in identifying an appropriate $\boldsymbol{R}_i$.

- Suppose we believe that, although there may be biological fluctuations over time, the observation times are sufficiently far apart that correlation due to within-subject sources among the Y_{ij} may be regarded as **negligible**.

In this case, it is reasonable to assume that $\text{var}(\boldsymbol{e}_{1i})$ is a **diagonal** matrix. If we furthermore believe that the magnitude of fluctuations is **similar** across time and units, we may represent this by the assumption that $\text{var}(e_{1ij}) = \sigma_1^2$, say, for all i and j, so that

$$\text{var}(\boldsymbol{e}_{1i}) = \sigma_1^2 \boldsymbol{I}_{n_i}.$$

The assumption that this is similar across units may be viewed as reflecting the belief that the e_{1ij} are **independent** of $\boldsymbol{\beta}_i$ and hence $\boldsymbol{b}_i$, which dictate how "large" the unit-specific trend is, so that the magnitude of fluctuations is unrelated to any unit-specific response characteristics.

- As we have discussed previously, it may be reasonable to assume that errors in measurement are **uncorrelated** over time; thus, taking $\text{var}(\boldsymbol{e}_{2i})$ to be a diagonal matrix would be appropriate.

  Suppose we also believe that errors committed by the measuring device are of similar magnitude regardless of the true size of the thing being measured, and are similar for all units (because the same device is used). This suggests that $\text{var}(e_{2ij}) = \sigma_2^2$, say, for all j, so that

  $$\text{var}(\boldsymbol{e}_{2i}) = \sigma_2^2 \boldsymbol{I}_{n_i}.$$

  Now the **true** size of the thing being measured at time t_{ij} is

  $$\beta_{0i} + \beta_{1i} t_{ij} + e_{1ij};$$

  i.e. the actual response uncontaminated by measurement error. Under this belief, it is reasonable to assume that the e_{2ij} are **independent** of $\boldsymbol{\beta}_i$ and thus $\boldsymbol{b}_i$.

- Putting this together, we would take

  $$\boldsymbol{R}_i = \text{var}(\boldsymbol{e}_i) = \text{var}(\boldsymbol{e}_{1i}) + \text{var}(\boldsymbol{e}_{2i}) = \sigma_1^2 \boldsymbol{I}_{n_i} + \sigma_2^2 \boldsymbol{I}_{n_i} = \sigma^2 \boldsymbol{I}_{n_i},$$

  where σ^2 is the aggregate variance reflecting variation due to both within-unit sources.

- The assumption that $\boldsymbol{e}_{1i}$ and $\boldsymbol{e}_{2i}$ are **independent** is **standard**, as is the assumption that $\boldsymbol{e}_{1i}$ and $\boldsymbol{e}_{2i}$ (and hence $\boldsymbol{e}_i$) are **independent** of $\boldsymbol{b}_i$. We say more about these assumptions shortly.

- We may think of other situations. For example, suppose that the response is something like **height**, which in all likelihood we can measure with very little if any error. Under this condition, we may effectively **eliminate** $\boldsymbol{e}_{2i}$ from the model and assume that $\boldsymbol{e}_i = \boldsymbol{e}_{1i}$; i.e. all within-unit variation is due to things like "fluctuations." In the model above, $\sigma^2 = \sigma_1^2$ would then represent the variance due to this sole source.

- Similarly, we may have a rather "noisy" measuring device such that, relative to errors in measurement, deviations due to within-unit subjects are virtually negligible. Under this condition, as long as we believe the times are far enough apart to render within-unit correlation negligible as well, we may as well take $\boldsymbol{e}_i = \boldsymbol{e}_{2i}$, in which case $\sigma^2 = \sigma_2^2$ in the above model represents solely measurement error variance.

- Now suppose that the times of observation are sufficiently close that correlation due to within-unit sources cannot be viewed as negligible. In this event, it would be unreasonable to take $\text{var}(\boldsymbol{e}_{1i})$ to be **diagonal**. It would instead be more realistic to adopt a model for $\text{var}(\boldsymbol{e}_{1i})$ that represents correlation that decays as observations become farther apart. For example, with equally-spaced observations and variance assumed constant as above, the AR(1) structure may be a suitable model; i.e.

$$\text{var}(\boldsymbol{e}_{1i}) = \sigma_1^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & \rho & 1 \end{pmatrix}.$$

  In general, maintaining the common variance assumption, we might entertain models $\text{var}(\boldsymbol{e}_{1i}) = \sigma_1^2 \boldsymbol{\Gamma}_i$, where $\boldsymbol{\Gamma}_i$ is a suitable $(n_i \times n_i)$ correlation matrix.

- In this case, with the same assumptions on measurement error and independence as above, we would instead have

$$\boldsymbol{R}_i = \text{var}(\boldsymbol{e}_i) = \sigma_1^2 \boldsymbol{\Gamma}_i + \sigma_2^2 \boldsymbol{I}_{n_i}. \tag{9.6}$$

  If measurement error were deemed negligible, this would be reduced to the assumption that

$$\boldsymbol{R}_i = \sigma^2 \boldsymbol{\Gamma}_i,$$

  where $\sigma^2 = \sigma_2^2$ represents variance due solely to within-unit fluctuations.

- We could also modify the above models to incorporate the possibility that, for example, one or both variances **changes** over time. In this situation, one could postulate a **heterogeneous** covariance model, as described in Chapter 4. I.e., if we believed fluctuation variances are still similar across subjects but change in magnitude over time, replace the assumption $\sigma_1^2 \boldsymbol{\Gamma}_i$ above by the heterogeneous version of the correlation matrix.

If we believe that there is a different variance at every time, this would make the most sense when all units are seen potentially at the same time points, as in the hip replacement study of the last chapter, so that there would be a finite number of variances to estimate. In this case, supposing there are n potential times at which units are seen, let $\text{var}(e_{1ij}) = \sigma^2_{1j}$ for the jth such time, $j = 1, \ldots, n$. Then for a unit seen at all n times, define

$$\boldsymbol{T}_i^{1/2} = \text{diag}(\sigma_{11}, \sigma_{12}, \ldots, \sigma_{1n}), \quad (n \times n),$$

where "diag" means a diagonal matrix with these values on the diagonal. We can then express the covariance matrix of the fluctuation deviation as

$$\text{var}(\boldsymbol{e}_{1i}) = \boldsymbol{T}_i^{1/2} \boldsymbol{\Gamma}_i \boldsymbol{T}_i^{1/2}$$

using the notation defined on page 45 in Chapter 3. For a unit with some time points missing, the considerations in the last chapter for specifying covariance matrices with unbalanced data would be used to write down the model for $\text{var}(\boldsymbol{e}_{1i})$ for each subject.

- Alternatively, it is conceivable that if there are several populations, $\boldsymbol{R}_i$ could be different for each. As an example, we could have

  $$\boldsymbol{R}_i = \sigma^2_G \boldsymbol{I}_{n_i} \quad \text{if } i \text{ is a girl}$$

  and $= Ri = \sigma^2_B \boldsymbol{I}_{n_i}$ if i is a boy, perhaps reflecting the belief that the magnitude of fluctuations is different for each gender.

- It should be clear that, in specifying the matrix $\boldsymbol{R}_i$, the analyst must consider carefully the features of the situation at hand in regard to within-unit sources of variation and correlation. Ideally, s/he would want to adopt a model that accurately characterizes the anticipated features.

- However, it turns out that, although not impossible, it may be difficult to fit a postulated model, particularly if it is rather complicated.

  For example, it is often problematic to fit models like (9.6) where **both** measurement error and "fluctuation" are assumed nonnegligible. This is often because there is not sufficient information to **identify** all the components of the model. A simplifying assumption that is thus often made is that one of the two sources tends to **dominate** the other. Under this assumption, modeling of $\boldsymbol{R}_i$ and fitting are simplified. The hope is that this may be a sufficiently good approximation to provide reliable inferences.

This sort of assumption is often made **unknowingly**; the analyst will choose a model for $\boldsymbol{R}_i$ that embodies certain assumptions and emphasizes one source or another by default without having thought about considerations like those above. In fact, the most common assumption is $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}_{n_i}$, where σ^2 is the same for all units and groups, is usually made in this way (and is the **default** in SAS `PROC MIXED`).

We discuss the consequences of a "wrong" model specification for $\boldsymbol{R}_i$ shortly.

- In general, $\boldsymbol{R}_i$ is a $(n_i \times n_i)$ matrix depending on a few variance and correlation parameters; e.g. σ^2 and ρ in the example above, chosen to at least approximate the anticipated features of within-unit sources of variation and correlation.

- If we just focus on the response for individual i at any time point t_{ij}, if we believe a **normal distribution** is a reasonable way to represent the population of responses we might see **on this individual** at t_{ij}, then it would make sense to assume that each e_{ij} were normally distributed. This of course implies that we assume

$$\boldsymbol{e}_i \sim \mathcal{N}_{n_i}(\boldsymbol{0}, \boldsymbol{R}_i).$$

AMONG-UNIT VARIATION: In the "population" model (9.5), the **random effects** $\boldsymbol{b}_i$ have mean $\boldsymbol{0}$ and represent variation resulting from the fact that individual units **differ**; i.e. exhibit **biological** or other variation. The model says that this variation **among individuals** manifests itself by causing the individual unit trajectories to be different (have different intercepts and slopes). Thus, $\mathrm{var}(\boldsymbol{b}_i)$ characterizes this variation.

- Intercepts and slopes may tend to be large or small **together**, so that children with steeper slopes tend to "start out" larger at age 0. Alternatively, large intercepts may tend to happen with small slopes and vice versa; perhaps children who "start out" smaller experience a steeper growth pattern to "catch up." In either case, this suggests that it would **not necessarily** be prudent to think of $\mathrm{var}(\boldsymbol{b}_i)$ as a **diagonal matrix**. Rather, we expect there to be some **correlation** between intercepts and slopes, the nature of this correlation depending on what is being studied.

- As noted above, we may believe that the populations of intercept/slopes for boys and girls have possibly different **means**, but that the variation in each population about the mean is similar. Formally, we can represent this by assuming that

$$\mathrm{var}(\boldsymbol{b}_i) = \boldsymbol{D}$$

for some covariance matrix $\boldsymbol{D}$ **regardless** of whether i is a boy or girl.

- Here, $\boldsymbol{D}$ is (2×2), and an **unstructured** model is really the only one that makes sense. In particular, writing

  $$\boldsymbol{D} = \begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{pmatrix},$$

  we have

  $$\mathrm{var}(\beta_{0i}) = \mathrm{var}(b_{0i}) = D_{11}, \quad \mathrm{var}(\beta_{1i}) = \mathrm{var}(b_{1i}) = D_{22}, \quad \mathrm{cov}(\beta_{0i}, \beta_{1i}) = \mathrm{cov}(b_{0i}, b_{1i}) = D_{12}.$$

  It should be clear that we would not expect $D_{12} = 0$ in general; e.g., steep slopes may be associated with "high" intercepts.

  It should also be clear that $D_{11} = D_{22}$ would be **unrealistic**. The **intercept** is on the same **scale** of measurement as the response, while the **slope** is on the scale "response scale per unit time." Thus, these parameters are representing variances that would be **expected** to be **different** because they correspond to phenomena that are on different scales.

- If we believed that these populations exhibit possibly different variation, we can represent this by assuming that

  $$\mathrm{var}(\boldsymbol{b}_i) = \boldsymbol{D}_B \text{ if } i \text{ is a boy}, \quad \mathrm{var}(\boldsymbol{b}_i) = \boldsymbol{D}_G \text{ if } i \text{ is a girl},$$

  where $\boldsymbol{D}_B$ and $\boldsymbol{D}_G$ are two (unstructured) covariance matrices.

- In either case, the assumption on $\mathrm{var}(\boldsymbol{b}_i)$ reflects **solely** the nature of variation at the level of the **population(s)** of units; that is, that caused **solely** by variation among units due to biology or other features. This is formally represented through the $\boldsymbol{b}_i$.

- It is often reasonable to assume that populations of intercepts and slopes are approximately **normally distributed**; e.g. this says that slopes vary **symmetrically** about the mean, some steeper, some shallower. Thus, a standard assumption is that the $\boldsymbol{b}_i$ have a **multivariate normal** distribution; e.g. in the case where the covariance matrix is assumed the same and equal to $\boldsymbol{D}$ regardless of gender, the assumption would be

  $$\boldsymbol{b}_i \sim \mathcal{N}_k(\boldsymbol{0}, \boldsymbol{D}),$$

  where k is the dimension of $\boldsymbol{b}_i$ ($k = 2$ here).

REMARKS:

- As noted previously, it is usually assumed that $\boldsymbol{e}_i$ and $\boldsymbol{b}_i$ are **independent**. This says that the magnitude of variation **within** a unit does not depend on the magnitude of $\boldsymbol{\beta}_i$ for that unit.

  As we have also discussed, if the device used to measure individual responses causes errors of similar magnitude all the time, and fluctuations are of similar magnitude regardless of the characteristics of the units, then this seems reasonable.

  However, if measurement errors tend to get larger as the response being measured gets larger, which is a characteristic of some measuring systems, then this may not be reasonable. In this case, we would expect the deviations in $\boldsymbol{e}_{2i}$ to be **related** to $\boldsymbol{Z}_i\boldsymbol{\beta}_i$ which dictates how large the responses on a particular unit are; we would also expect them to be related to the deviations in $\boldsymbol{e}_{1i}$.

  Similarly, if the magnitude of fluctuations is related to inherent unit characteristics (e.g., "high" units tend to have larger fluctuations), the assumption would also be violated.

- We will assume for now that this assumption is reasonable, and take $\boldsymbol{b}_i$ and $\boldsymbol{e}_i$ to be **independent**, as is customary. Later, we will discuss situations where this is definitely unreasonable in more detail.

- We have also noted that specification of the within-units covariance matrix $\boldsymbol{R}_i$ to reflect reality is desirable. However, computational issues and a tendency to not consider the issue carefully can lead to choice of an unrealistic model.

- As we will see in a moment, the specifications on $\mathrm{var}(\boldsymbol{b}_i)$ and $\mathrm{var}(\boldsymbol{e}_i)$ **combine** to produce an **overall model** for $\mathrm{var}(\boldsymbol{\epsilon}_i)$ that describes the aggregate effects of both sources of variation. The hope is that this model is **rich** and **flexible** enough that it can still represent the true pattern of overall variation even if one or both components are **incorrectly** modeled.

  If interest focuses only on $\boldsymbol{\beta}$, this may be adequate. However, if there is interest in how units **vary in the population**, represented by $\mathrm{var}(\boldsymbol{b}_i)$, it seems clear that getting this model correct is **essential**. We will say more later.

SUMMARY: We now summarize the model suggested by these considerations. The model may be thought of as a **two-stage hierarchy**: For $i = 1, \ldots, m$,

Stage 1 – individual

$$\boldsymbol{Y}_i = \boldsymbol{Z}_i \boldsymbol{\beta}_i + \boldsymbol{e}_i \quad (n_i \times 1), \quad \boldsymbol{e}_i \sim \mathcal{N}_{n_i}(\boldsymbol{0}, \boldsymbol{R}_i) \tag{9.7}$$

This is like a "regression model" for the ith unit, with "design matrix" $\boldsymbol{Z}_i$ and $(k \times 1)$ "regression parameter" $\boldsymbol{\beta}_i$.

Stage 2 – population

$$\boldsymbol{\beta}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i \quad (k \times 1), \quad \boldsymbol{b}_i \sim \mathcal{N}_k(\boldsymbol{0}, \boldsymbol{D}). \tag{9.8}$$

Here, we have taken $\text{var}(\boldsymbol{b}_i) = \boldsymbol{D}$ to be the **same** for all i, and we will continue to do so for definiteness in our subsequent development. However, this could be relaxed as described above, and the features of the model we point out shortly would still be valid. The matrix $\boldsymbol{A}_i$ summarizes information like group membership, allowing the **mean** of $\boldsymbol{\beta}_i$ to be different for different groups.

Variation in the model is explicitly acknowledged to come from **two** sources:

- Due to features **within** units, represented through the covariance matrix $\boldsymbol{R}_i$.
- Due to biological variation **among** units, represented to the covariance matrix $\boldsymbol{D}$.
- This is in marked contrast to the models of the previous chapter. These models required the analyst to think of a **single** covariance matrix for a data vector, representing the aggregate effect of **both sources**. The models that are typically used tend to focus on the time-ordered aspect.

IMPLICATION: We now see the contrast with the models of the last chapter more directly. Suppose that we **combine** two parts of the model into a single representation by substituting the expression for $\boldsymbol{\beta}_i$ in (9.8) into (9.7); i.e.

$$\boldsymbol{Y}_i = \boldsymbol{Z}_i(\boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i) + \boldsymbol{e}_i = (\boldsymbol{Z}_i \boldsymbol{A}_i)\boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i.$$

- Suppose first that there is only one group, so that $\boldsymbol{A}_i = \boldsymbol{I}_k$. Then we see that the model implied is

  $$\boldsymbol{Y}_i = \boldsymbol{Z}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i.$$

  Note that we can write this in a more familiar form by letting $\boldsymbol{X}_i = \boldsymbol{Z}_i$ and $\boldsymbol{\epsilon}_i = \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i$. With these identifications, we have

  $$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, m.$$

This has exactly the form of the regression models of the previous chapter!

- The difference is that, here, the way we arrived at this model requires that the error vector $\boldsymbol{\epsilon}_i$ have the **particular** form above. Note that this implies that, using the independence of $\boldsymbol{b}_i$ and $\boldsymbol{e}_i$ (and taking $\text{var}(\boldsymbol{b}_i) = \boldsymbol{D}$ for definiteness),

$$\text{var}(\boldsymbol{\epsilon}_i) = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{R}_i = \boldsymbol{\Sigma}_i. \tag{9.9}$$

Thus, the model implied by thinking in two stages implies that the covariance matrix of a data vector is the sum of **two** pieces representing the **separate** effects of among-and within-unit variation.

- If there is more than one group, the same interpretation holds. Suppose $\boldsymbol{\beta}$ is $(p \times 1)$; $p = 4$ in the dental example. With $\boldsymbol{\beta}_i$ $(k \times 1)$, then $\boldsymbol{A}_i$ a $(k \times p)$ matrix; $k = 2$ in the dental example. Then we see that the model implied is

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i,$$

where $\boldsymbol{X}_i = \boldsymbol{Z}_i \boldsymbol{A}_i$. As above, $\text{var}(\boldsymbol{\epsilon}_i)$ is as in (9.9). In the dental example, note that for boys

$$\boldsymbol{X}_i = \boldsymbol{Z}_i \boldsymbol{A}_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \\ 1 & t_{in_i} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & t_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & t_{in_i} \end{pmatrix}$$

and similarly for girls,

$$\boldsymbol{X}_i = \boldsymbol{Z}_i \boldsymbol{A}_i = \begin{pmatrix} 1 & t_{i1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{in_i} & 0 & 0 \end{pmatrix}.$$

Compare these with (8.9); they are **the same**.

RESULT: By thinking about individual trajectories, we see that we ultimately arrive at a regression model that is of the **same form** as those in the last chapter.

- The **similarity** is that the **mean** of a data vector is of the same **linear** form; i.e.

$$E(\boldsymbol{Y}_i) = \boldsymbol{X}_i \boldsymbol{\beta},$$

where the form of the matrices $\boldsymbol{X}_i$ is dictated by the thinking above ($\boldsymbol{X}_i = \boldsymbol{Z}_i \boldsymbol{A}_i$).

The **critical difference** is that the covariance matrix of a data vector has the very specific form (9.9) that explicitly acknowledges **both** sources of variation and allows them to be thought about **separately**. Further features of note:

- The model does **not** allow the covariance matrix of a data vector to be the **same** for all units in general. The only way that this matrix may be of the same form for all units is $\text{var}(\boldsymbol{b}_i)$ and $\text{var}(\boldsymbol{e}_i)$ are the same for all units and the data are **balanced** (more on this shortly).
- The covariance matrix **depends** on the times of observation through the matrix $\boldsymbol{Z}_i$. Thus, if different units are seen at different times, this information is **automatically** incorporated into the model.
- Recall that we have noted that we expect observations on the same unit to be correlated **even if** the repeated observations are taken very far apart in time; this is due to the simple fact that they are from the **same unit**. Note that the implied form of the covariance matrix (9.9) accommodates this naturally. Even if $\boldsymbol{R}_i = \sigma^2\boldsymbol{I}$, say, which implies that we believe there is no correlation due to within-unit sources, the entire matrix $\boldsymbol{\Sigma}_i$ is still **not** diagonal. Rather, it will be nondiagonal because $\boldsymbol{D}$ is not diagonal in general. Thus, the model offers a natural way to represent correlation among observations on the same unit that arises simply because they are on the same unit and thus "more alike" than those compared across units.
- In this model, $\boldsymbol{\Sigma}_i$ depends on a finite set of parameters. For example, if $\boldsymbol{R}_i = \sigma^2\boldsymbol{I}_{n_i}$, then $\boldsymbol{\Sigma}_i$ depends on σ^2 and the **distinct** elements of the matrix $\boldsymbol{D}$. We say **distinct** because, as $\boldsymbol{D}$ is a covariance matrix, it is **symmetric**, so contains the same off-diagonal elements more than once; e.g. if
$$\boldsymbol{D} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix},$$
then $\boldsymbol{D}$ depends on the three distinct values D_{11}, D_{12}, and D_{22}, since $D_{12} = D_{21}$ by symmetry.

We may in fact say even more. If we believe that both $\boldsymbol{b}_i$ and $\boldsymbol{e}_i$ are both well-represented by multivariate normal distributions and are independent, then, using results in Chapter 4, we may conclude that

$$\boldsymbol{Y}_i \sim \mathcal{N}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}_i), \quad i = 1, \ldots, m \tag{9.10}$$
$$\boldsymbol{X}_i = \boldsymbol{Z}_i\boldsymbol{A}_i, \quad \boldsymbol{\Sigma}_i = \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i' + \boldsymbol{R}_i.$$

- As with the models of the previous chapter, if the units are completely unrelated, then it is reasonable to assume that the $\boldsymbol{Y}_i$ are **independent** random vectors, each multivariate normal with the particular mean and covariance structure given above.

TERMINOLOGY: These models are known as **random coefficient** models because they rely on thinking of individual-specific **regression parameters**, or **coefficients** of time, as being **random**, each representing a draw from a population.

- The above reasoning is extended easily to the case where units come from more than two groups; for example, for the dialyzer data, where the relationship between transmembrane pressure ("time") and ultrafiltration rate (response) was observed on dialyzers from 3 centers. We would thus think of each dialyzer having its own straight line relationship, with its own intercept and slope ($k = 2$). The vector $\boldsymbol{\beta}$ would represent the **mean** intercept and slope for each center stacked together, so would have $p = 6$ elements.
- The reasoning is extended easily to the case where the "regression model" for an individual unit is something other than a **straight line**; e.g. suppose a quadratic function is a better model (recall the hip replacement data)

  $$Y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} + \beta_{2i} t_{ij}^2 + e_{ij}.$$

  In this case, $\boldsymbol{\beta}_i$ has $k = 3$ elements.
- All of these models are a particular case of the more general class of **linear mixed effects models** we will describe in the next chapter.

9.3 Inference on regression and covariance parameters

Because this way of thinking leads ultimately to the model given in (9.10), the methods of **maximum likelihood** and **restricted maximum likelihood** may be used to estimate the parameters that characterize "mean" and "variation," namely $\boldsymbol{\beta}$, the distinct elements of $\boldsymbol{D}$, and the parameters that make up $\boldsymbol{R}_i$. That is, the methods described in sections 8.5 and 8.6 may be used exactly as described. The same considerations apply:

- The **generalized least squares** estimator for $\boldsymbol{\beta}$ and its large sample approximate sampling distribution will have the same form, with $\boldsymbol{X}_i$ and $\boldsymbol{\Sigma}_i$ defined as in (9.10).
- Questions of interest may be written in the identical fashion, and estimation of approximate standard errors, Wald tests, likelihood ratio tests for nested models, and so on may be carried out in the same way. We will discuss the **formulation** and **interpretation** of questions of interest under this model momentarily.

- Information criteria may be used to compare non-nested models.

See these sections for descriptions, which go through unchanged for the model (9.10).

QUESTIONS OF INTEREST: Because of the way we motivated the random coefficient model, questions of interest may be thought of in different ways. For definiteness, again consider the situation of the dental study data. A **vague** statement of the main question of interest is: “Is the rate of change of distance as children age different for boys and girls?”

Both here and in the previous chapter, we end up with a model that says that the **mean** of all possible Y_{ij} values we might see at a particular age t_{ij} for girls is

$$E(Y_{ij}) = \beta_{0,G} + \beta_{1,G} t_{ij},$$

and similarly for boys. How we arrive at the model involved different thinking, however.

- In the previous chapters, we always thought in terms of how the **means** at each time were related, averaged across all units at each time point. In this way of thinking, we write down the model above immediately, and $\beta_{1,G}$ and $\beta_{1,B}$ have the interpretation as the parameters that describe the relationship of the **mean responses** over time; that is, the slope of the (assumed straight line) relationship among means at different times t_{ij}.
- From the motivation for the random coefficient model, we think in terms of individual trajectories and their “typical” features. In this way of thinking, $\beta_{1,G}$ and $\beta_{1,B}$ have the interpretation as the **means** of the populations of child-specific slopes for all possible girls and boys, respectively.

Since the model we end up with is the **same**, **either** interpretation is valid. The result is that we may think of the vague question of interest more formally in two ways, and both are correct. If we consider testing

$$H_0 : \beta_{1,G} - \beta_{1,B} = 0 \text{ vs. } H_1 : \beta_{1,G} - \beta_{1,B} \neq 0,$$

we may interpret this as saying either of the following:

1. Does the rate of change in mean response over time differ between girls and boys?
2. Is the “typical” value of the slope of the individual straight lines for girls different from the “typical” value of the slope of the individual straight lines for boys?

THE "TYPICAL" PROFILE VS THE "TYPICAL" RATE OF CHANGE: This fuss over how to state the vague question of interest and interpret this statement may seem to be overblown. However, it has some important practical consequences.

- Depending on the subject matter, one interpretation may make more sense than another. The **process** occurring over time may be something that is naturally thought of as happening **within** a unit, such as **growth**. Under these circumstances, an investigator may find it easier to think in terms of the random coefficient model, which says that each child has his/her own individual trajectory with his/her own rate of change (slope). Then the question is naturally one about the comparison of "typical" (mean) slopes.
- In other contexts, investigators may find it easier to think in terms of the "typical" response **profile**; i.e. how the means across all units over time change. This might be true if the ultimate goal is to make public policy recommendations. If the response is score on an achievement test administered to each of m children each year for 5 years in two different curricula, the investigator is interested in how the means over children in each group change over time; he would like to claim that the average score for one curriculum got better faster than the other. His thinking will tend to focus on how change happens over time to children as a group (means) rather than on "typical" change over time for children.

The distinction in interpretation is quite a subtle one, and most people find it difficult to grasp at first. As we have seen, **either** interpretation makes sense for our model.

- As we will see later, this is because the model both for mean response as a function of time and the individual trajectories is **linear** in the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_i$.
- When this model is **not** linear, we will see that the interpretation gets more difficult.

ALTERNATIVE FITTING METHOD: A natural inclination when thinking about random coefficient models is to exploit the fact that the model says that each unit has its own trajectory and hence own "regression model" with unit-specific "regression parameter" $\boldsymbol{\beta}_i$, where the $\boldsymbol{\beta}_i$ come from a population with mean ("typical value") $\boldsymbol{\beta}$. (We discuss one population here, but the following reasoning applies to more than one.) This suggests that if we want to learn about $\boldsymbol{\beta}$, a one way to do it would be to **estimate** each $\boldsymbol{\beta}_i$ from each unit **separately**, and then **combine** the results to estimate $\boldsymbol{\beta}$; e.g. estimate $\boldsymbol{\beta}$ as the **sample mean** of the individual unit estimates of $\boldsymbol{\beta}_i$.

- Such an approach represents an alternative to fitting the full model by ML or REML as discussed above, and is often called a **two-stage** estimation method. This is because fitting happens in two stages.

- (1) Estimate each $\boldsymbol{\beta}_i$ separately from the data on unit i only; e.g. if we believe $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}_{n_i}$ for each i, then we might estimate $\boldsymbol{\beta}_i$ by usual least squares applied to the data from unit i. Call these estimates $\widehat{\boldsymbol{\beta}}_i$.

- (2) This distills the data $\boldsymbol{Y}_i$ on each individual down to new "data" $\widehat{\boldsymbol{\beta}}_i$. This suggests using the new "data" as the basis for inference. For example, a natural approach would be to average the $\widehat{\boldsymbol{\beta}}_i$ across all i to estimate $\boldsymbol{\beta}$; e.g. if there is only one group, estimate $\boldsymbol{\beta}$ as

  $$m^{-1} \sum_{i=1}^{m} \widehat{\boldsymbol{\beta}}_i.$$

  If there are several groups, do this on a group by group basis, e.g. average the estimates from boys and girls separately.

- To compare groups, compare these sample averages of estimates across groups by using standard statistical methods, e.g. apply an analysis of variance to the slope estimates to compare the mean slope.

This sounds appealing, but it isn't quite right.

- The new "data," the individual estimates $\widehat{\boldsymbol{\beta}}_i$, are not exactly the "data" we'd like. The ideal for learning about $\boldsymbol{\beta}$ would be to average the **true** $\boldsymbol{\beta}_i$ across units. Of course, we don't know these and the best we can do is estimate them by $\widehat{\boldsymbol{\beta}}_i$. But this introduces additional **uncertainty** that the above procedure does not take into account.

- For example, if the n_i are very different across units, with some units having lots of measurements and others only a few, then for some i, $\widehat{\boldsymbol{\beta}}_i$ will be a better estimate of the true $\boldsymbol{\beta}_i$ than for others. Treating them all on equal footing as "data" is thus obviously not appropriate.

- Thus, simply averaging the $\widehat{\boldsymbol{\beta}}_i$ as if they were the true $\boldsymbol{\beta}_i$ can be misleading.

It turns out that if one wants to use individual estimates as "data," one must instead take a **weighted** average of the $\widehat{\boldsymbol{\beta}}_i$ in an appropriate way to take these issues into account. This kind of approach is discussed in Davidian and Giltinan (1995).

Historically, the use of two-stage methods was suggested quite a long time ago, in part because it made intuitive sense. A fundamental paper advocating two-stage methods is Rowell and Walters (1976). Other references to two-stage methods include Gumpertz and Pantula (1989) and Davidian and Giltinan (1995). Because the methods of ML and REML are straightforward to implement with available software, we do not consider two-stage methods further here.

SPECIAL CASE – BALANCED DATA: Recall in the last chapter we noted an interesting curiosity for the dental data, which are **balanced**. When we assumed that the covariance matrix of a data vector, $\boldsymbol{\Sigma}_i$ (which is actually the same for all i with balanced data) had the **compound symmetry** structure, we saw that the generalized least squares estimator for $\boldsymbol{\beta}$ reduced to the ordinary least squares estimator $\widehat{\boldsymbol{\beta}}_{OLS}$ treating all data as if they were independent. That is, the GLS estimator

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{m} \boldsymbol{X}_i' \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{X}_i \right)^{-1} \sum_{i=1}^{m} \boldsymbol{X}_i' \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{Y}_i \tag{9.11}$$

with $\boldsymbol{\Sigma}$ having the **compound symmetry** structure had the same value as the OLS estimator

$$\widehat{\boldsymbol{\beta}}_{OLS} = \left(\sum_{i=1}^{m} \boldsymbol{X}_i' \boldsymbol{X}_i \right)^{-1} \sum_{i=1}^{m} \boldsymbol{X}_i' \boldsymbol{Y}_i.$$

It turns out that this is a special instance of a more general result. The general result says:

- For the **random coefficient model**, if (i) the data are **balanced**, with all units seen at the **same** n times, so that the design matrix $\boldsymbol{Z}_i$ of time points is the **same** for all units i, **and** (ii) $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}_n$, then then the generalized least squares estimator is numerically equivalent to the OLS estimator!

- To show this is a **nasty** but not impossible exercise in matrix algebra. Under conditions (i) and (ii), $\boldsymbol{\Sigma}_i$ reduces to the **same** matrix for each i:

  $$\boldsymbol{\Sigma}_i = \boldsymbol{Z}\boldsymbol{D}\boldsymbol{Z}' + \sigma^2 \boldsymbol{I}_n.$$

  Substitute this expression for $\widehat{\boldsymbol{\Sigma}}$ in (9.11) for each i (even if $\boldsymbol{D}$ and σ^2 are replaced by estimates, the form is the same). Fancy footwork with matrix inversion formulæ like those in Chapter 2 may then be used to show the equivalence. Those with strong stomachs might want to try it!

The **compound symmetry** assumption for $\boldsymbol{\Sigma}$ **directly** in these circumstances is just a special case of the particular covariance structure $\boldsymbol{\Sigma}_i = \boldsymbol{ZDZ}' + \sigma^2 \boldsymbol{I}_n$ for balanced data. To see this, consider a simple model with one group, so that

$$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_j + e_{ij},$$
$$\text{var}(\boldsymbol{b}_i) = \boldsymbol{D} = \begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{pmatrix}, \quad \text{var}(\boldsymbol{e}_i) = \sigma^2 \boldsymbol{I}_n.$$

- It is straightforward to verify that (try it!)

$$\text{var}(Y_{ij}) = D_{11} + D_{22}t_j^2 + 2D_{12}t_j + \sigma^2, \quad \text{cov}(Y_{ij}, Y_{ik}) = D_{11} + D_{22}t_j t_k + D_{12}(t_j + t_k).$$

- Note that if $D_{22} = 0$ and $D_{12} = 0$, then these reduce to

$$\text{var}(Y_{ij}) = D_{11} + \sigma^2, \quad \text{cov}(Y_{ij}, Y_{ik}) = D_{11},$$

  which is the compound symmetry model!

NEED FOR COVARIANCE STRUCTURE: As we have stressed before, just because the GLS estimator is numerically identical to the OLS estimator under these circumstances is no reason to disregard the need to characterize the covariance structure of a data vector correctly!

- The approximate covariance matrix of the GLS estimator, $\widehat{\boldsymbol{V}}_\beta$, **depends** on the form of $\boldsymbol{\Sigma}_i$, even if the estimator $\widehat{\boldsymbol{\beta}}$ doesn't!

9.4 Inference on individuals

The random coefficient model is intuitively appealing – it comes from thinking first about individuals and their own unique trajectories, and then about the population of individuals (in terms of the parameters that characterize these trajectories). Thinking this way leads to a model for the mean and covariance of a data vector that has a specific form; in particular, the covariance matrix of data vector is represented explicitly as the sum of 2 terms, incorporating separately the impact of 2 sources of variation, within- and among-units. This makes it easier for the data analyst:

- The sources of variation may be thought of separately. Thus, for example, a model $\boldsymbol{R}_i$ that best captures the variation due to the nature of data collection on an individual unit may be entertained separately from having to think about biological variation ($\boldsymbol{D}$). In the modeling approach of the last chapter, this had to be done all at once.

The model has still another advantage. It is sometimes the case that investigators may wish not only to learn about the **population(s)** of units through things such as the "typical" (mean) slope values and how they compare across populations. Particularly in medical and educational studies, the investigators may wish to understand the change in the response over time for **specific subjects**.

- In a study of AIDS patients, with response "viral load," measuring "amount" of virus in the system, investigators may wish to characterize the trajectory of viral load for particular patients in order to aid in decisions about their future care.
- In educational studies, where response is some measure of "achievement," investigators may wish to characterize the progress of individual children in order to place them in the most suitable learning environment.

If we think in terms of the random coefficient model, then, interest focuses on the **subject-specific** parameters $\boldsymbol{\beta}_i$ describing the trajectories of individual subjects. In particular, for individual subjects, the investigators are interested in "estimating" $\boldsymbol{\beta}_i$ for specific subjects based on the data.

- One way to do this would be just to use estimates based on treating each subject as a separate regression problem – one could get $\widehat{\boldsymbol{\beta}}_i$ from each subject's data separately.
- However, if the numbers of observations on each i is not too large, these estimates will probably not be very good.
- Moreover, this does not take into account (nor does it take advantage of) the fact that we have data from an entire sample of **similar** subjects from the same population(s). Intuition suggests that we could stand to gain something from acknowledging that we believe this!

We will take up this issue in the next chapter, when we discuss the general **linear mixed effects model**, of which the random coefficient model is a special case.

- Note immediately, however, that the models we have talked about in this course up to now (Chapters 4–7) do not even explicitly acknowledge individual trajectories!

9.5 Discussion

"POPULATION-AVERAGED" VS. "SUBJECT-SPECIFIC": We have seen that the random coefficient model arises from thinking about the longitudinal data situation in an alternative way. Rather than thinking in terms of the **mean responses** at each time point and how they are related, we think of **individual trajectories** and then the **means** of individual-specific parameters that characterize these trajectories (e.g. mean of the slopes in the population of subjects).

- The first approach, which was used in Chapter 7, is often called a **population-averaged** approach for this reason – the focus of modeling is on the **averages** (**means**) across the **population** of units at each time point, and how these averages are related over time.
- The current approach is often called a **subject-specific** approach – the focus of modeling is on individual units.
- In the case where the models considered are **linear**, the two perspectives ultimately lead to the **same** type of model for the mean, so that either interpretation is valid.
- The subject-specific, random coefficient approach has the additional feature that it "automatically" leads to a particular assumption about the structure of the covariance matrix of a data vector, which naturally acknowledges within- and among-unit variation separately. In contrast, the population-averaged approach forces the data analyst to model this covariance, thinking about the two sources of variation **together**. As a result, the subject-specific approach of the random coefficient model, and, more generally, the **linear mixed effects models** we will consider in the next chapter, has become incredibly popular.

ALTERNATIVE TERMINOLOGY: The random coefficient model, allowing for the possibility of different **groups**, is sometimes referred to as a **growth curve** model in the statistical and subject-matter literature.

CHOICE OF COVARIANCE STRUCTURE: We have noted that the possibilities are quite broad for modeling covariance structure within the random coefficient model framework.

- One may in principle take the covariance matrix $\boldsymbol{R}_i$, corresponding to **within-unit** variation, to be one of a variety of structures according to knowledge of the data collection process.
- If the main source of within-unit variation is measurement error, or if it is instead fluctuation but observations are far apart in time taking $\boldsymbol{R}_i$ diagonal may be reasonable.

- One may in principle take the covariance matrix var($\boldsymbol{b}_i$), characterizing variation **among units** (through how the parameters in the individual trajectories vary) to be the same for all groups or different, depending on the belief about the pattern of variation for each group.

- The most commonly-used form of the random coefficient model is that where

$$\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}_{n_i}, \quad \text{var}(\boldsymbol{b}_i) = \boldsymbol{D} = \text{ same for all groups.}$$

  Often this structure is suitable; e.g. units tend to vary similarly for each group, although the means may be different (same $\boldsymbol{D}$ is reasonable). This same kind of assumption (means differ, variance the same) is standard in usual analysis of variance models and methods. This model is considered extensively and almost exclusively in much of the literature. It is certainly possible to relax these assumptions; for example, we discussed the possibility of taking $\boldsymbol{D}$ to be different for each gender group in the dental data example.

- One pitfall of trying to get too fancy with modeling of $\boldsymbol{R}_i$ and var($\boldsymbol{b}_i$) is that it is quite likely that one will end up with a model that is **too complicated** to be sorted out given the data at hand. This problem of **identifiability** is mentioned in the next section.

- Thus, many people are willing to risk the possibility that they may incorrectly specify $\boldsymbol{R}_i$ and/or $\boldsymbol{D}$ by, for example, assuming that thevar($\boldsymbol{b}_i$) = $\boldsymbol{D}$ is common to all groups when it may not be. The form of the model

$$\boldsymbol{\Sigma}_i = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{R}_i$$

  is sufficiently general that, even if the two components $\boldsymbol{D}$ and $\boldsymbol{R}_i$ are not **exactly** correctly chosen, the resulting $\boldsymbol{\Sigma}_i$ matrix will differ very little from that one would obtain if they were. Thus, if one's main interest is in estimating $\boldsymbol{\beta}$ and tests about it, this may be okay.

- However, if interest is focused on var($\boldsymbol{b}_i$) and $\boldsymbol{R}_i$ themselves, then obviously one would want to investigate all possibilities. Thus, in the first example of section 9.7, we illustrate how both the commonly-used specification and fancier ones may be implemented in SAS. However, be aware that fitting very fancy models may lead to difficulties and "over-fitting." To read more about the possibilities, see *SAS System for Mixed Models* (1996, chapter 8) and Vonesh and Chinchilli (1997, section 6.3).

9.6 Basic `PROC MIXED` sytnax

We are now in a position to explain fully exactly how `PROC MIXED` is set up. In the most general case of a random coefficient model, we may write the model as

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{e}_i.$$

In fact, just as we did in the previous chapter, we may present this mode in a streamlined form by "stacking" the contributions from each unit. In particular, Define

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \\ \vdots \\ \boldsymbol{Y}_m \end{pmatrix}, \quad \boldsymbol{e} = \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_m \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_m \end{pmatrix}, \quad \boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{R}_m \end{pmatrix},$$

$$\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_m \end{pmatrix}, \quad \boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{Z}_m \end{pmatrix}, \quad \widetilde{\boldsymbol{D}} = \begin{pmatrix} \boldsymbol{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{D} \end{pmatrix},$$

where $\widetilde{\boldsymbol{D}}$ here has been displayed in the case where $\text{var}(\boldsymbol{b}_i) = \boldsymbol{D}$ for all units but could be modified if, say, girls and boys had different matrices $\boldsymbol{D}_G$ and $\boldsymbol{D}_B$. We may then write the model concisely as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{e}, \quad \text{var}(\boldsymbol{Y}) = \boldsymbol{Z}\widetilde{\boldsymbol{D}}\boldsymbol{Z}' + \boldsymbol{R} \tag{9.12}$$

(verify). This type of concise expression is used in the documentation, except that SAS refers to $\widetilde{\boldsymbol{D}}$ as $\boldsymbol{G}$.

We have already seen that the `model` statement is the mechanism by which the analyst may specify the form the **mean vector**, denoted $\boldsymbol{X}_i\boldsymbol{\beta}$ for unit i or $\boldsymbol{X}\boldsymbol{\beta}$ for all units, stacked. We have used the `repeated` statement to specify the overall covariance matrix.

- In the context of a model of the above form, however, the `repeated` statement is used to specify the **within-unit** covariance model $\boldsymbol{R}_i$ or, equivalently, $\boldsymbol{R}$ above.
- An additional statement, the `random` statement, is used to specify the assumption on $\text{var}(\boldsymbol{b}_i)$ ($\widetilde{\boldsymbol{D}}$).

We will see specific examples in the next section.

For now, we offer a summary of the basic syntax for quick reference.

```
proc mixed data=dataset method= (ML,REML);
class  classification variables;
model response =  columns of X / solution;
random columns of Z / type=  subject=  group=  ;
repeated / type=  subject=  group=  ;
run;
```

`proc mixed` statement

- `method=REML` is the default; no `method=` required in this case

`model` statement

- **columns of $\boldsymbol{X}$** are variables (`class` or continuous) corresponding to variables associated with fixed effects $\boldsymbol{\beta}$
- Intercept is assumed unless `noint` option after slash
- `solution` is an option

`random` statement

- Describes the matrix $\widetilde{\boldsymbol{D}} = \text{var}(\boldsymbol{b})$ (i.e. the matrices var$(\boldsymbol{b}_i)$ making up the blocks of $\widetilde{\boldsymbol{D}}$
- **columns of $\boldsymbol{Z}$** are variables (`class` or continuous), i.e. variables associated with random effects $\boldsymbol{b}$
- `subject=` tells `mixed` what `class` variable denotes the grouping determining the **units**
- `type=` allows choice of matrix (e.g. `un`, unstructured)
- `group=` allows $\boldsymbol{D}$ to be different according to this `class` variable (e.g. dental study, boys, girls)

`repeated` statement

- Describes the matrix $\boldsymbol{R} = \text{var}(\boldsymbol{e})$ (i.e. the matrices $\boldsymbol{R}_i = \text{var}(\boldsymbol{e}_i)$
- If $\text{var}(\boldsymbol{e}_i) = \sigma^2 \boldsymbol{I}_{n_i}$ same for all i `repeated` statement is *NOT* needed
- `subject=` tells `mixed` what `class` variable denotes the grouping determining the units
- `type=` allows choice other than diagonal (e.g. `ar(1), cs, etc.`
- `group=` allows $\boldsymbol{R}_i$ to be different depending on group membership (e.g. dental study, $\text{var}(\boldsymbol{e}_i) = \sigma^2_G$ girls, $\text{var}(\boldsymbol{e}_i) = \sigma^2_B$ boys)

We may now observe that, in the previous chapter, to implement a general linear regression model using `proc mixed` with the `repeated` statement, we simply made a correspondence between the model of form (9.12) with **no** random effects $\boldsymbol{b}$, which looks like

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e},$$

and the model in that chapter of the form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}_i.$$

From purely **operational** point of view (but **not** an **interpretation** point of view), the models have the same structure – a mean plus a deviation with components of length n_i, each of which has a covariance matrix. Thus, purely to specify these covariance matrices for the second model, the `repeated` statement can be used.

See the SAS documentation for `PROC MIXED` for much more detail on the use of these statements and available options.

9.7 Implementation with SAS

We illustrate how to carry out analyses based on random coefficient models for two examples we have already considered:

1. The dental study data
2. The ultrafiltration data

For each data set, we consider different random coefficient models and address questions of interest such as whether the mean slope differs across groups (gender or center). As discussed in the last section, we use SAS `PROC MIXED` with the `random` statement to impose the random coefficient model structure – this statement allows the user to specify $\text{var}(\boldsymbol{b}_i)$. If there is no `repeated` statement, it is assumed that $\text{var}(\boldsymbol{e}_i) = \sigma^2 \boldsymbol{I}_{n_i}$ (see the last section). Otherwise, if a `random` and `repeated` statement appear simultaneously, the `repeated` statement sets up some other model for $\text{var}(\boldsymbol{e}_i) = \boldsymbol{R}_i$.

WARNING – LACK OF IDENTIFIABILITY: It is important to use `PROC MIXED` with version 6.12 or higher of SAS; here, we use version 8.2. Even with this improved version, as well as with programs in other software packages that are designed to fit these models, things may not always go as planned. It is important to keep in mind that the models are being fit via numerical algorithms that are used to maximize the likelihood or restricted likelihood. It is possible to specify a model with $\text{var}(\boldsymbol{b}_i)$ and $\text{var}(\boldsymbol{e}_i)$ sufficiently complex that it is **too complicated** to be fitted given the information available in the data. That is, one may choose these models in such a way that there are too many parameters, more than are required to give an adequate characterization of the true covariance structure. Such a model is said to be **over-identified** or **unidentifiable**. The result of specifying such models is that the numerical algorithms will either fail to find a solution (converge) or will lead to a solution that is **nonsensical**). Thus, one **pitfall** to be aware of when fitting these models and more generally those of the next chapter is the possibility of getting "carried away" in choosing the structure for $\boldsymbol{R}_i$, making it too complicated and leading to an **unidentifiable** model. If `PROC MIXED` fails to converge for a particular model choice, then the analyst may have to consider whether the implied model for $\boldsymbol{\Sigma}_i$ is "too rich" for the problem and adopt simpler choices (at the risk of being "wrong").

EXAMPLE 1 – DENTAL STUDY DATA:

- For illustration purposes only, we fit the random coefficient model assuming that the mean intercept and slope differ for the two genders. Note that when fitting a random coefficient model, it is natural to think in terms of the parameterization of the model that contains intercept and slope explicitly rather than their difference:

  $$Y_{ij} = \beta_{0i} + \beta_{1i}t_{ij} + e_{ij},$$

  $$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{b}_i, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_{0,G} \\ \beta_{1,G} \end{pmatrix} \text{ girls}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_{0,B} \\ \beta_{1,B} \end{pmatrix} \text{ boys}.$$

  We consider this parameterization in our fitting.

- For fitting this model, we illustrate how to instruct `PROC MIXED` to fit models for a number of different assumptions on the matrices $\boldsymbol{R}_i$ and $\text{var}(\boldsymbol{b}_i)$. These are:

  (i) $\boldsymbol{R}_i = \sigma^2\boldsymbol{I}$, $\boldsymbol{D}$ same for both genders. This is the most common specification. Recall this implies a belief that within-child sources of correlation are negligible ($\boldsymbol{R}_i$ **diagonal**) and among-child variation is similar in each group. The parameter σ^2 may be interpreted as the aggregate variance due to within-child "fluctuations" in distance and measurement error.

  (ii) $\boldsymbol{R}_i = \sigma^2_G\boldsymbol{I}$ if i is a girl and $\boldsymbol{R}_i = \sigma^2_B\boldsymbol{I}$ if i is a boy, $\boldsymbol{D}$ same for both genders. This allows for the possibility that within-child variation might be different for the different genders (due to measurement error and fluctuation).

  (iii) $\boldsymbol{R}_i$ is the AR(1) covariance matrix, same for both genders, and $\boldsymbol{D}$ is the same for both genders. This choice of $\boldsymbol{R}_i$ allows for the possibility of nonnegligible within-child correlation.

  (iv) $\boldsymbol{R}_i = \sigma^2_G\boldsymbol{I}$ if i is a girl and $\boldsymbol{R}_i = \sigma^2_B\boldsymbol{I}$ if i is a boy, and $\text{var}(\boldsymbol{b}_i) = \boldsymbol{D}_G$ if i is a girl and $= \boldsymbol{D}_B$ if a boy. This allows for the possibility that within-child variation might be different for the different genders **and** the possibility that variability in intercepts and slopes is different. This essentially amounts to fitting two separate models, one for each gender!

  (v) $\boldsymbol{R}_i$ is the sum of two components: an AR(1) covariance matrix (corresponding to the fluctuations, allowing within-child correlation) and $\sigma^2_2\boldsymbol{I}$, which now corresponds to the measurement error component (assumed common). $\boldsymbol{D}$ is the same for both genders. Specifically, we have

  $$\boldsymbol{R}_i = \sigma^2_1\boldsymbol{\Gamma} + \sigma^2_2\boldsymbol{I},$$

  where $\boldsymbol{\Gamma}$ is the (4×4) AR(1) correlation matrix. To fit this model, we use of the `local` option of the `repeated` statement, which adds the matrix $\sigma^2_2\boldsymbol{I}$ to the requested AR(1) matrix.

PROGRAM:

```
/******************************************************************

  CHAPTER 9, EXAMPLE 1

  Analysis of the dental study data by fitting a random coefficient
  model in time using PROC MIXED.

  -  the repeated measurement factor is age (time)

  -  there is one "treatment" factor, gender

  The model for each child is assumed to be a straight line.
  The intercepts and slopes may have different means depending on
  gender, with the same covariance matrix D for each gender.

  We use the RANDOM and REPEATED statements to fit models that
  make several different assumptions about the forms of the matrices
  Ri and D.

******************************************************************/

options ls=80 ps=59 nodate; run;

/*****************************************************************

  Read in the data set (See Example 1 of Chapter 4)

******************************************************************/

data dent1; infile 'dental.dat';
  input obsno child age distance gender;
run;

/******************************************************************

  Use PROC MIXED to fit the random coefficient model via the
  RANDOM statement.  For all of the fits, we use usual normal
  ML rather than REML (the default).

  In all cases, we use the usual parameterization for the mean
  model.

  The SOLUTION option in the MODEL statement requests that the
  estimates of the regression parameters be printed.

  The G and GCORR options in the RANDOM statement asks that the
  D matrix and the corresponding correlation matrix it implies
  be printed.  The V and VCORR options ask that the overall
  Sigma matrix be printed (for the first subject or particular
  subjects).

  To fit a random coefficient model, we must specify that both
  intercept and slope are random in the RANDOM statement.

  If no REPEATED statement appears, then PROC MIXED assumes that
  Ri =  sigma^2*I.  Otherwise, we use a REPEATED statement to set
  a structure for Ri with the TYPE = option.

******************************************************************/

*  MODEL (i);
*  Ri = diagonal with constant variance sigma^2 same in both genders;
*  No REPEATED statement necessary to fit this Ri (default);
*  D = (2x2) unstructured matrix same for both genders;
*  Specified in the RANDOM statement;

title 'RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD';
title2 'COVARIANCE MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER';
title3 'SAME D MATRIX FOR BOTH GENDERS';
proc mixed method=ml data=dent1;
  class gender child;
  model distance = gender gender*age / noint solution;
  random intercept age / type=un subject=child g gcorr v vcorr;
  estimate 'diff in mean slope' gender 0 0 gender*age 1 -1;
  contrast 'overall gender diff' gender 1 -1, gender*age 1 -1 /chisq;
run;

*  MODEL (ii);
*  Fit the same model but with a separate diagonal Ri matrix for;
*  each gender.  Thus, there are 2 separate variances sigma^2_(G and B);
*  D still = (2x2) unstructured matrix same for both genders;
*  Specified in the RANDOM statement;

title 'RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD';
title2 'COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER';
```

```
title3 'SAME D MATRIX FOR BOTH GENDERS';
proc mixed  method=ml data=dent1;
  class  child gender;
  model distance = gender gender*age / noint solution;
  repeated / group=gender subject=child;
  random intercept age / type=un subject=child g gcorr v vcorr;
  estimate 'diff in mean slope' gender 0 0 gender*age 1 -1;
  contrast 'overall gender diff' gender 1 -1, gender*age 1 -1 /chisq;
run;

*  MODEL (iii);
*  Ri is AR(1) with the same variance and rho value for each gender;
*  Specified in the REPEATED statement;
*  D still = (2x2) unstructured matrix same for both genders;
*  Specified in the RANDOM statement;

title 'RANDOM COEFFICIENT MODEL WITH AR(1) WITHIN-CHILD';
title2 'CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER';
title3 'SAME D MATRIX FOR BOTH GENDERS';
proc mixed method=ml data=dent1;
  class gender child ;
  model distance = gender gender*age / noint solution ;
  random intercept age / type=un subject=child g gcorr v vcorr;
  repeated / type=ar(1) subject=child rcorr;
  estimate 'diff in mean slope' gender 0 0 gender*age 1 -1;
  contrast 'overall gender diff' gender 1 -1, gender*age 1 -1 /chisq;
run;

*  MODEL (iv);
*  Fit the same model but with a separate diagonal Ri matrix for;
*  each gender.  Thus, there are 2 separate variances sigma^2_(G and B);
*  D still = (2x2) unstructured matrix differs across genders;
*  Specified in the RANDOM statement by the GROUP=GENDER option;

title 'RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD';
title2 'COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER';
title3 'DIFFERENT D MATRIX FOR BOTH GENDERS';
proc mixed  method=ml data=dent1;
  class  child gender;
  model distance = gender gender*age / noint solution;
  repeated / group=gender subject=child;
  random intercept age / type=un group=gender subject=child g gcorr v vcorr;
  estimate 'diff in mean slope' gender 0 0 gender*age 1 -1;
  contrast 'overall gender diff' gender 1 -1, gender*age 1 -1 /chisq;
run;

*  MODEL (v)
*  Ri is the sum of two components, an AR(1) component for fluctuations;
*  and a diagonal component with variance sigma^2 common to both genders;
*  The LOCAL option adds the diagonal component to the AR(1) structure;
*  specified in the REPEATED statement;
*  D still = (2x2) unstructured matrix same for both genders;
*  Specified in the RANDOM statement;

title 'RANDOM COEFFICIENT MODEL WITH AR(1) + COMMON MEAS ERROR WITHIN-CHILD';
title2 'CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER';
title3 'SAME D MATRIX FOR BOTH GENDERS';
proc mixed method=ml data=dent1;
  class gender child ;
  model distance = gender gender*age / noint solution ;
  random intercept age / type=un subject=child g gcorr v vcorr;
  repeated / type=ar(1) local subject=child rcorr;
  estimate 'diff in mean slope' gender 0 0 gender*age 1 -1;
  contrast 'overall gender diff' gender 1 -1, gender*age 1 -1 /chisq;
run;
```

OUTPUT: Following the output, we comment on a few aspects of the output.

```
      RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD              1
 COVARIANCE MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                SAME D MATRIX FOR BOTH GENDERS

                     The Mixed Procedure

                      Model Information

    Data Set                     WORK.DENT1
    Dependent Variable           distance
    Covariance Structure         Unstructured
    Subject Effect               child
    Estimation Method            ML
    Residual Variance Method     Profile
    Fixed Effects SE Method      Model-Based
    Degrees of Freedom Method    Containment

                   Class Level Information

      Class     Levels    Values

      gender         2    0 1
      child         27    1 2 3 4 5 6 7 8 9 10 11 12 13
                          14 15 16 17 18 19 20 21 22 23
                          24 25 26 27

                         Dimensions

             Covariance Parameters             4
             Columns in X                      4
             Columns in Z Per Subject          2
             Subjects                         27
             Max Obs Per Subject               4

                   Number of Observations

         Number of Observations Read             108
         Number of Observations Used             108
         Number of Observations Not Used           0

                      Iteration History

 Iteration    Evaluations        -2 Log Like       Criterion

         0              1       478.24175986
         1              1       427.80595080      0.00000000

                  Convergence criteria met.

     RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD              2
 COVARIANCE MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                SAME D MATRIX FOR BOTH GENDERS

                     The Mixed Procedure

                     Estimated G Matrix

       Row    Effect       child         Col1        Col2

         1    Intercept     1          4.5569     -0.1983
         2    age           1         -0.1983     0.02376

                Estimated G Correlation Matrix

       Row    Effect       child         Col1        Col2

         1    Intercept     1          1.0000     -0.6025
         2    age           1         -0.6025      1.0000

                Estimated V Matrix for child 1

      Row        Col1        Col2        Col3        Col4

        1      4.6216      2.8891      2.8727      2.8563
        2      2.8891      4.6839      3.0464      3.1251
        3      2.8727      3.0464      4.9363      3.3938
        4      2.8563      3.1251      3.3938      5.3788

          Estimated V Correlation Matrix for child 1

      Row        Col1        Col2        Col3        Col4

        1      1.0000      0.6209      0.6014      0.5729
        2      0.6209      1.0000      0.6335      0.6226
        3      0.6014      0.6335      1.0000      0.6586
        4      0.5729      0.6226      0.6586      1.0000
```

```
                         Covariance Parameter Estimates

                        Cov Parm     Subject    Estimate

                        UN(1,1)      child        4.5569
                        UN(2,1)      child       -0.1983
                        UN(2,2)      child       0.02376
                        Residual                  1.7162

                                Fit Statistics

                     -2 Log Likelihood              427.8
                     AIC (smaller is better)        443.8
                     AICC (smaller is better)       445.3
                     BIC (smaller is better)        454.2

             RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD             3
         COVARIANCE MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                         SAME D MATRIX FOR BOTH GENDERS

                              The Mixed Procedure

                        Null Model Likelihood Ratio Test

                          DF    Chi-Square      Pr > ChiSq

                           3         50.44          <.0001

                          Solution for Fixed Effects

                                      Standard
  Effect        gender    Estimate       Error      DF    t Value    Pr > |t|

  gender        0          17.3727      1.1820      54      14.70      <.0001
  gender        1          16.3406      0.9801      54      16.67      <.0001
  age*gender    0           0.4795     0.09980      54       4.80      <.0001
  age*gender    1           0.7844     0.08275      54       9.48      <.0001

                         Type 3 Tests of Fixed Effects

                               Num     Den
                Effect          DF      DF    F Value    Pr > F

                gender           2      54     247.00    <.0001
                age*gender       2      54      56.46    <.0001

                                   Estimates

                                     Standard
   Label                 Estimate       Error      DF    t Value    Pr > |t|

   diff in mean slope     -0.3048      0.1296      54      -2.35      0.0224

                                   Contrasts

                       Num    Den
Label                   DF     DF    Chi-Square   F Value     Pr > ChiSq   Pr > F

overall gender diff      2     54         14.19      7.10         0.0008   0.0018


              RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD             4
       COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                         SAME D MATRIX FOR BOTH GENDERS

                              The Mixed Procedure

                               Model Information

             Data Set                     WORK.DENT1
             Dependent Variable           distance
             Covariance Structures        Unstructured, Variance
                                          Components
             Subject Effects              child, child
             Group Effect                 gender
             Estimation Method            ML
             Residual Variance Method     None
             Fixed Effects SE Method      Model-Based
             Degrees of Freedom Method    Containment

                            Class Level Information

               Class     Levels    Values

               child         27    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27
               gender         2    0 1

                                  Dimensions
```

```
                    Covariance Parameters             5
                    Columns in X                      4
                    Columns in Z Per Subject          2
                    Subjects                         27
                    Max Obs Per Subject               4

                          Number of Observations

                Number of Observations Read              108
                Number of Observations Used              108
                Number of Observations Not Used            0

                             Iteration History

        Iteration    Evaluations        -2 Log Like       Criterion

                0              1       478.24175986
                1              2       418.92503842      1.16632499
                2              1       416.18869903      1.23326209
                3              1       407.89638533      0.01954268
                4              2       406.88264563      0.00645800
                5              1       406.10632159      0.00056866
                6              1       406.04318997      0.00000764
                7              1       406.04238894      0.00000000

            RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD              5
     COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                         Convergence criteria met.

                            Estimated G Matrix

              Row    Effect       child        Col1        Col2

                1    Intercept    1          3.1978     -0.1103
                2    age          1         -0.1103     0.01976

                       Estimated G Correlation Matrix

              Row    Effect       child        Col1        Col2

                1    Intercept    1          1.0000     -0.4388
                2    age          1         -0.4388      1.0000

                       Estimated V Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      3.1426      2.7933      2.8889      2.9845
               2      2.7933      3.4128      3.1426      3.3172
               3      2.8889      3.1426      3.8411      3.6499
               4      2.9845      3.3172      3.6499      4.4275

                 Estimated V Correlation Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      1.0000      0.8529      0.8315      0.8001
               2      0.8529      1.0000      0.8680      0.8534
               3      0.8315      0.8680      1.0000      0.8851
               4      0.8001      0.8534      0.8851      1.0000

                       Covariance Parameter Estimates

                Cov Parm     Subject    Group       Estimate

                UN(1,1)      child                    3.1978
                UN(2,1)      child                   -0.1103
                UN(2,2)      child                   0.01976
                Residual     child      gender 0      0.4449
                Residual     child      gender 1      2.6294

           RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD              6
     COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                              Fit Statistics

                   -2 Log Likelihood             406.0
                   AIC (smaller is better)       424.0
                   AICC (smaller is better)      425.9
                   BIC (smaller is better)       435.7

                      Null Model Likelihood Ratio Test
```

```
                          DF    Chi-Square      Pr > ChiSq

                           4         72.20          <.0001

                          Solution for Fixed Effects

                                      Standard
  Effect        gender    Estimate       Error     DF     t Value     Pr > |t|

  gender        0          17.3727      0.7386     54       23.52       <.0001
  gender        1          16.3406      1.1114     54       14.70       <.0001
  age*gender    0           0.4795     0.06180     54        7.76       <.0001
  age*gender    1           0.7844     0.09722     54        8.07       <.0001

                         Type 3 Tests of Fixed Effects

                               Num     Den
                Effect          DF      DF    F Value    Pr > F

                gender           2      54     384.72    <.0001
                age*gender       2      54      62.66    <.0001

                                   Estimates

                                     Standard
   Label                 Estimate       Error      DF    t Value    Pr > |t|

   diff in mean slope     -0.3048      0.1152      54      -2.65      0.0106

                                   Contrasts

                       Num    Den
Label                   DF     DF   Chi-Square   F Value     Pr > ChiSq   Pr > F

overall gender diff      2     54        14.32      7.16         0.0008   0.0017

                RANDOM COEFFICIENT MODEL WITH AR(1) WITHIN-CHILD               7
         CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                         SAME D MATRIX FOR BOTH GENDERS

                              The Mixed Procedure

                               Model Information

             Data Set                    WORK.DENT1
             Dependent Variable          distance
             Covariance Structures       Unstructured,
                                         Autoregressive
             Subject Effects             child, child
             Estimation Method           ML
             Residual Variance Method    Profile
             Fixed Effects SE Method     Model-Based
             Degrees of Freedom Method   Containment

                            Class Level Information

               Class     Levels    Values

               gender         2    0 1
               child         27    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27

                                  Dimensions

                      Covariance Parameters             5
                      Columns in X                      4
                      Columns in Z Per Subject          2
                      Subjects                         27
                      Max Obs Per Subject               4

                            Number of Observations

                  Number of Observations Read             108
                  Number of Observations Used             108
                  Number of Observations Not Used           0

                               Iteration History

          Iteration    Evaluations        -2 Log Like       Criterion

                  0              1       478.24175986
                  1              2       424.08934703      0.00028001
                  2              1       424.05684775      0.00000096
                  3              1       424.05673965      0.00000000

                           Convergence criteria met.

               RANDOM COEFFICIENT MODEL WITH AR(1) WITHIN-CHILD               8
```

```
       CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                 Estimated R Correlation Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      1.0000     -0.4680      0.2190     -0.1025
               2     -0.4680      1.0000     -0.4680      0.2190
               3      0.2190     -0.4680      1.0000     -0.4680
               4     -0.1025      0.2190     -0.4680      1.0000

                             Estimated G Matrix

              Row    Effect       child        Col1        Col2

                1    Intercept     1        10.1459     -0.7198
                2    age           1        -0.7198     0.07508

                       Estimated G Correlation Matrix

              Row    Effect       child        Col1        Col2

                1    Intercept     1         1.0000     -0.8248
                2    age           1        -0.8248      1.0000

                       Estimated V Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      4.6275      2.6363      3.2182      2.5959
               2      2.6363      4.4510      2.7601      3.6423
               3      3.2182      2.7601      4.8751      3.4846
               4      2.5959      3.6423      3.4846      5.8999

                 Estimated V Correlation Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      1.0000      0.5809      0.6776      0.4968
               2      0.5809      1.0000      0.5925      0.7108
               3      0.6776      0.5925      1.0000      0.6497
               4      0.4968      0.7108      0.6497      1.0000

                       Covariance Parameter Estimates

                      Cov Parm     Subject    Estimate

                      UN(1,1)      child       10.1459
                      UN(2,1)      child       -0.7198
                      UN(2,2)      child       0.07508
             RANDOM COEFFICIENT MODEL WITH AR(1) WITHIN-CHILD               9
       CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                       Covariance Parameter Estimates

                      Cov Parm     Subject    Estimate

                      AR(1)        child       -0.4680
                      Residual                  1.1940

                              Fit Statistics

                   -2 Log Likelihood             424.1
                   AIC (smaller is better)       442.1
                   AICC (smaller is better)      443.9
                   BIC (smaller is better)       453.7

                      Null Model Likelihood Ratio Test

                        DF    Chi-Square      Pr > ChiSq

                         4         54.19          <.0001

                        Solution for Fixed Effects

                                    Standard
Effect        gender    Estimate       Error      DF    t Value    Pr > |t|

gender        0          17.4166      1.1586      54      15.03      <.0001
gender        1          16.1544      0.9607      54      16.82      <.0001
age*gender    0           0.4757      0.1010      54       4.71      <.0001
age*gender    1           0.7978     0.08374      54       9.53      <.0001
```

```
                        Type 3 Tests of Fixed Effects

                              Num     Den
                Effect         DF      DF    F Value    Pr > F

                gender          2      54     254.37    <.0001
                age*gender      2      54      56.48    <.0001

                                  Estimates

                                        Standard
   Label                  Estimate         Error      DF    t Value    Pr > |t|

   diff in mean slope      -0.3220        0.1312      54      -2.45      0.0174
               RANDOM COEFFICIENT MODEL WITH AR(1) WITHIN-CHILD              10
         CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                        SAME D MATRIX FOR BOTH GENDERS

                             The Mixed Procedure

                                  Contrasts

                       Num     Den
Label                   DF      DF    Chi-Square   F Value     Pr > ChiSq   Pr > F

overall gender diff      2      54         13.46      6.73         0.0012   0.0025
             RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD             11
       COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                     DIFFERENT D MATRIX FOR BOTH GENDERS

                             The Mixed Procedure

                              Model Information

             Data Set                        WORK.DENT1
             Dependent Variable              distance
             Covariance Structures           Unstructured, Variance
                                             Components
             Subject Effects                 child, child
             Group Effects                   gender, gender
             Estimation Method               ML
             Residual Variance Method        None
             Fixed Effects SE Method         Model-Based
             Degrees of Freedom Method       Containment

                           Class Level Information

               Class     Levels    Values

               child         27    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27
               gender         2    0 1

                                 Dimensions

                      Covariance Parameters             8
                      Columns in X                      4
                      Columns in Z Per Subject          4
                      Subjects                         27
                      Max Obs Per Subject               4

                           Number of Observations

                  Number of Observations Read             108
                  Number of Observations Used             108
                  Number of Observations Not Used           0

                              Iteration History

          Iteration    Evaluations        -2 Log Like       Criterion

                  0              1       478.24175986
                  1              1       405.11800674      0.00000000

                          Convergence criteria met.
             RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD             12
       COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                     DIFFERENT D MATRIX FOR BOTH GENDERS

                             The Mixed Procedure

                              Estimated G Matrix

  Row    Effect      child    gender         Col1        Col2        Col3        Col4

    1    Intercept    1       0            2.9716    -0.07539
    2    age          1       0          -0.07539     0.02151
```

```
   3   Intercept    1      1                                 5.6468    -0.2827
   4   age          1      1                                -0.2827    0.02530

                       Estimated G Correlation Matrix

 Row   Effect      child   gender       Col1        Col2        Col3        Col4

   1   Intercept    1      0          1.0000     -0.2982
   2   age          1      0         -0.2982      1.0000
   3   Intercept    1      1                                 1.0000     -0.7480
   4   age          1      1                                -0.7480      1.0000

                       Estimated V Matrix for child 1

              Row        Col1        Col2        Col3        Col4

                1      3.5889      3.3357      3.5292      3.7226
                2      3.3357      4.0618      3.8947      4.1742
                3      3.5292      3.8947      4.7069      4.6258
                4      3.7226      4.1742      4.6258      5.5240

                 Estimated V Correlation Matrix for child 1

              Row        Col1        Col2        Col3        Col4

                1      1.0000      0.8737      0.8587      0.8361
                2      0.8737      1.0000      0.8907      0.8812
                3      0.8587      0.8907      1.0000      0.9072
                4      0.8361      0.8812      0.9072      1.0000

                       Covariance Parameter Estimates

                 Cov Parm     Subject    Group       Estimate

                 UN(1,1)      child      gender 0      2.9716
                 UN(2,1)      child      gender 0    -0.07539
                 UN(2,2)      child      gender 0     0.02151
                 UN(1,1)      child      gender 1      5.6468
                 UN(2,1)      child      gender 1     -0.2827
                 UN(2,2)      child      gender 1     0.02530
                 Residual     child      gender 0      0.4466
                 Residual     child      gender 1      2.5891

            RANDOM COEFFICIENT MODEL WITH DIAGONAL WITHIN-CHILD            13
      COVARIANCE MATRIX WITH SEPARATE CONSTANT VARIANCE FOR EACH GENDER
                     DIFFERENT D MATRIX FOR BOTH GENDERS

                             The Mixed Procedure

                               Fit Statistics

                    -2 Log Likelihood              405.1
                    AIC (smaller is better)        429.1
                    AICC (smaller is better)       432.4
                    BIC (smaller is better)        444.7

                      Null Model Likelihood Ratio Test

                        DF    Chi-Square      Pr > ChiSq

                         7         73.12          <.0001

                        Solution for Fixed Effects

                                          Standard
 Effect        gender    Estimate        Error      DF    t Value    Pr > |t|

 gender        0          17.3727       0.7252      25      23.96      <.0001
 gender        1          16.3406       1.1715      25      13.95      <.0001
 age*gender    0           0.4795      0.06313      25       7.60      <.0001
 age*gender    1           0.7844      0.09835      25       7.98      <.0001

                       Type 3 Tests of Fixed Effects

                             Num     Den
              Effect          DF      DF    F Value    Pr > F

              gender           2      25     384.22    <.0001
              age*gender       2      25      60.65    <.0001

                                 Estimates

                                         Standard
  Label                   Estimate        Error      DF    t Value    Pr > |t|

  diff in mean slope       -0.3048       0.1169      25      -2.61      0.0151

                                 Contrasts
```

```
                       Num    Den
Label                   DF     DF    Chi-Square   F Value     Pr > ChiSq   Pr > F

overall gender diff      2     25         14.12      7.06         0.0009   0.0037

     RANDOM COEFFICIENT MODEL WITH AR(1) + COMMON MEAS ERROR WITHIN-CHILD    14
         CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                         SAME D MATRIX FOR BOTH GENDERS

                              The Mixed Procedure

                               Model Information

             Data Set                     WORK.DENT1
             Dependent Variable           distance
             Covariance Structures        Unstructured,
                                          Autoregressive
             Subject Effects              child, child
             Estimation Method            ML
             Residual Variance Method     Profile
             Fixed Effects SE Method      Model-Based
             Degrees of Freedom Method    Containment

                            Class Level Information

               Class     Levels    Values

               gender         2    0 1
               child         27    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27

                                  Dimensions

                      Covariance Parameters             6
                      Columns in X                      4
                      Columns in Z Per Subject          2
                      Subjects                         27
                      Max Obs Per Subject               4

                            Number of Observations

                  Number of Observations Read             108
                  Number of Observations Used             108
                  Number of Observations Not Used           0

                               Iteration History

          Iteration    Evaluations        -2 Log Like       Criterion

                  0              1       478.24175986
                  1              2       428.22548286     24.55088017
                  2              2       427.26075815      1.09477678
                  3              2       426.51452533      1.16919129
                  4              2       425.99015592      0.08543213
                  5              2       424.91951841      0.01458002
                  6              2       424.32018203      0.00323017
                  7              3       424.01683319       .
                  8              1       423.99457950      0.00007763

     RANDOM COEFFICIENT MODEL WITH AR(1) + COMMON MEAS ERROR WITHIN-CHILD    15
         CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                         SAME D MATRIX FOR BOTH GENDERS

                              The Mixed Procedure

                               Iteration History

          Iteration    Evaluations        -2 Log Like       Criterion

                  9              1       423.99420143      0.00000054
                 10              2       423.99415208      0.00000007
                 11              2       423.99414400      0.00000000

                           Convergence criteria met.


                   Estimated R Correlation Matrix for child 1

               Row         Col1         Col2         Col3         Col4

                 1       1.0000      -0.2256       0.2241      -0.2227
                 2      -0.2256       1.0000      -0.2256       0.2241
                 3       0.2241      -0.2256       1.0000      -0.2256
                 4      -0.2227       0.2241      -0.2256       1.0000

                               Estimated G Matrix

                Row    Effect       child         Col1         Col2
```

```
                1    Intercept     1         6.9045     -0.4333
                2    age           1        -0.4333     0.04828

                       Estimated G Correlation Matrix

              Row    Effect       child        Col1        Col2

                1    Intercept     1         1.0000     -0.7505
                2    age           1        -0.7505      1.0000

                       Estimated V Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      4.5375      2.6344      3.2041      2.4504
               2      2.6344      4.5423      2.8323      3.5951
               3      3.2041      2.8323      4.9333      3.4165
               4      2.4504      3.5951      3.4165      5.7106

   RANDOM COEFFICIENT MODEL WITH AR(1) + COMMON MEAS ERROR WITHIN-CHILD    16
       CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                 Estimated V Correlation Matrix for child 1

             Row        Col1        Col2        Col3        Col4

               1      1.0000      0.5803      0.6772      0.4814
               2      0.5803      1.0000      0.5983      0.7059
               3      0.6772      0.5983      1.0000      0.6437
               4      0.4814      0.7059      0.6437      1.0000

                       Covariance Parameter Estimates

                      Cov Parm     Subject    Estimate

                      UN(1,1)      child        6.9045
                      UN(2,1)      child       -0.4333
                      UN(2,2)      child       0.04828
                      Variance     child        0.3351
                      AR(1)        child       -0.9935
                      Residual                  1.1408

                              Fit Statistics

                   -2 Log Likelihood              424.0
                   AIC (smaller is better)        444.0
                   AICC (smaller is better)       446.3
                   BIC (smaller is better)        457.0

                      Null Model Likelihood Ratio Test

                        DF    Chi-Square      Pr > ChiSq

                         5         54.25          <.0001

                        Solution for Fixed Effects

                                    Standard
Effect        gender    Estimate       Error      DF    t Value    Pr > |t|

gender        0          17.4148      1.1651      54      14.95      <.0001
gender        1          16.1917      0.9661      54      16.76      <.0001
age*gender    0           0.4757      0.1010      54       4.71      <.0001
age*gender    1           0.7979     0.08376      54       9.53      <.0001
   RANDOM COEFFICIENT MODEL WITH AR(1) + COMMON MEAS ERROR WITHIN-CHILD    17
       CORRELATION MATRIX WITH CONSTANT VARIANCE SAME FOR EACH GENDER
                       SAME D MATRIX FOR BOTH GENDERS

                            The Mixed Procedure

                       Type 3 Tests of Fixed Effects

                             Num     Den
              Effect           DF      DF    F Value    Pr > F

              gender            2      54     252.17    <.0001
              age*gender        2      54      56.46    <.0001

                                 Estimates

                                   Standard
 Label                 Estimate       Error      DF    t Value    Pr > |t|

 diff in mean slope     -0.3222      0.1312      54      -2.45      0.0173

                                 Contrasts
```

```
                          Num     Den
Label                      DF      DF    Chi-Square    F Value      Pr > ChiSq    Pr > F

overall gender diff         2      54         13.97       6.99          0.0009    0.0020
```

INTERPRETATION:

- For each assumed model, the output shows the estimates of $\boldsymbol{D}$ (or different such matrices where appropriate), the estimates of parameters making up $\boldsymbol{R}_i$, and, as usual, the estimates of $\boldsymbol{\beta}$. For the fit of model (i), the estimate of the assumed common $\boldsymbol{D}$ is (`Estimated G Matrix`) and the implied correlation matrix (`Estimated G Correlation Matrix`) are

$$\begin{pmatrix} 4.5569 & -0.1983 \\ -0.1983 & 0.02376 \end{pmatrix}, \quad \begin{pmatrix} 1.0000 & -0.6025 \\ -0.6025 & 1.0000 \end{pmatrix},$$

respectively. The estimate of σ^2 in the assumed model $\boldsymbol{R}_i = \sigma^2\boldsymbol{I}$ is in the `Covariance Parameter Estimates` table (along with the distinct elements of $\boldsymbol{D}$ repeated) and is equal to 1.716 (`Residual`). Recall that these are **balanced** data; thus, under this assumption, the matrix $\boldsymbol{\Sigma}_i$ is the **same** for all children. The estimate of $\boldsymbol{\Sigma}_i$ implied by the above estimates and the associated correlation matrix are given in the tables `Estimated V Matrix for CHILD 1` and `Estimated V Correlation Matrix for CHILD 1` (see the output, page 1 and 2).

For the other models (ii) – (v), the estimates of the components of the overall covariance structure are given in a similar fashion. For model (ii), the estimates of $\boldsymbol{D}$ and its implied correlation matrix appear on page 5 of the output. Here, we assume that the within-child variance is different depending on gender; from the table `Covariance Parameter Estimates`, the estimates are given as $\hat{\sigma}^2_G = 0.445$ and $\hat{\sigma}^2_B = 2.629$. These estimates are quite different. The implied matrix $\boldsymbol{\Sigma}_i$ is now different for different i; in particular, it will be the same for all boys and the same for all girls. The `v` and `vcorr` options cause `PROC MIXED` to print the estimate of $\boldsymbol{\Sigma}_i$ for the first child, so the estimates of `Estimated V Matrix for CHILD 1` and `Estimated V Correlation Matrix for CHILD 1` correspond to the estimate for girls.

For the fit of model (iii), where a common AR(1) structure is assumed for both boys and girls, the estimates of ρ and σ^2 may be found on page 9–10 of the output in the table `Covariance Parameter Estimates` as -0.468 and 1.194, respectively.

For model (iv), where a different $\boldsymbol{D}$ matrix and $\boldsymbol{R}_i$ matrix as in model (ii) are assumed for each gender, SAS prints the estimates of the two matrices $\boldsymbol{D}_G$ and $\boldsymbol{D}_B$ in the `Estimated G Matrix` together on page 12; that for girls is

$$\begin{pmatrix} 2.9716 & -0.0754 \\ -0.0754 & 0.0215 \end{pmatrix}.$$

The corresponding correlation matrices are printed in `Estimated G Correlation Matrix`. Again, the implied $\mathbf{\Sigma}_i$ matrices will differ for boys and girls; those for the first girl are printed on page 12.

For model (v), which included two components for $\boldsymbol{R}_i$, results begin on page 14 of the output. In the `Covariance Parameter Estimates` table, `Variance` is generated by the `local` option and refers to the estimate of σ_2^2. `Residual` refers to the common variance σ_1^2 that appears as part of the structure requested in `type=`. `AR(1)` refers to the estimate of ρ. Note that the estimated value is -0.99, which is virtually 1! The estimate has wandered off toward the "boundary" of what its possible values are. Note that the overall covariance model is very "rich." This is typical behavior under these conditions and probably reflects that this model is too fancy to be well-identified.

- Note in cases (i), (ii), and (iv) that the estimates of $\boldsymbol{\beta}$ found in the `Solution for Fixed Effects` are identical and are equal to the ordinary least squares estimator. This reflects the argument given in section 9.3. Of course, the estimated standard errors are different for the different fits, reflecting the different assumptions about $\mathbf{\Sigma}_i$ that go into forming $\widehat{\boldsymbol{V}}_\beta$. For (iii) and (v), where the AR(1) matrix is involved so that $\boldsymbol{R}_i$ does not have a form like $sigma^2\boldsymbol{I}$ for all units, this does not hold.

- For all analyses, the Wald test of different slopes carried out by the `estimate` statement gives a significant result at level $\alpha = 0.05$. Also obtained is a Wald test for the "overall difference" between genders – the $\boldsymbol{L}$ matrix for this contrast is

$$\boldsymbol{L} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix};$$

thus, we are testing whether the mean intercepts and slopes are the same for each gender **simultaneously**. Regardless of the assumption on $\mathbf{\Sigma}_i$, the evidence supporting rejection of this null hypothesis seems very strong.

- Inspection of the AIC and BIC values for these fits on pages 2, 6, 9, and 13 of the output shows that model (ii), where a different within-child variance is assumed for each gender and $\boldsymbol{D}$ the same seem preferable among the four models considered. The AIC and BIC values for this model are 424.0 and 435.7, respectively. Comparing the values to those for the general regression models considered in the analysis of these data in section 8.8 reveals that these AIC and BIC values seem comparable to those for the preferred model in that section, where $\boldsymbol{\Sigma}_i$ was modeled as following a different compound symmetry structure for boys and girls. Thus, among all models considered for these data so far, either of these seems plausible. Model (ii) here may be more pleasing to many analysts, because it considers the two sources of variation explicitly. The key element seems to be allowing the within-child variance to be different for the two genders; allowing $\boldsymbol{D}$ to differ as well in model (iv) offered no improvement in fit. Inspection of the original data plot reveals the potential source of this result. Note that 2 of the boys, and one especially, have trajectories that seem to "bounce around" much more than those of the other children. From above, the estimate of variance for boys, σ^2_B, was much larger than that for girls, σ^2_G. Otherwise, the trajectories seem similarly spread out across girls and boys, supporting the choice of common $\boldsymbol{D}$. Being able to model the covariance structure in terms of the two sources of variation explicitly makes this clear, allowing a pleasing interpretation of how the overall covariance structure differs. Such an interpretation is more difficult with the model of section 8.8.

EXAMPLE 2 – DIALYZER DATA: In the following program, we consider the issue of whether the mean slope of a trajectory differs across the centers.

- The "full" model is that assuming that each dialyzer has its own straight line trajectory with its own intercept and slope. Then, each center has its own mean intercept and slope. We assume a common var($\boldsymbol{b}_i$) = $\boldsymbol{D}$ and a common diagonal within-unit covariance matrix $\boldsymbol{R}_i = \sigma^2 \boldsymbol{I}$ for all centers. Other specifications could be investigated to see if they provide a better fit.

- The model is

  $$Y_{ij} = \beta_{0i} + \beta_{i1} t_{ij} + e_{ij},$$

  $$\boldsymbol{\beta}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{02} \\ \beta_{12} \\ \beta_{03} \\ \beta_{13} \end{pmatrix}, \quad \boldsymbol{b}_i \sim \mathcal{N}_2(\boldsymbol{0}, \boldsymbol{D}).$$

  where $\beta_{0\ell}$, $\beta_{1\ell}$ are the mean intercept and slope for the ℓth center, $\ell = 1, 2, 3$. $\boldsymbol{A}_i$ is the appropriate matrix of 0's and 1's that "picks off" the correct elements of $\boldsymbol{\beta}$ for the i dialyzer; e.g. if i is from center 1, then

  $$\boldsymbol{A}_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

  We fit this model by ML and REML.

- We also consider the reduced model where the slopes are the **same** for each center (with different intercepts). Thus, for this model

  $$\boldsymbol{\beta} = \begin{pmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{03} \\ \beta_{1} \end{pmatrix},$$

  where β_1 is the common slope. Thus, $\boldsymbol{A}_i$ would be the (2×4) matrix to "pick off" the right intercept and β_1 for the ith center; e.g. for i from center 1,

  $$\boldsymbol{A}_i = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

  We fit this model by ML so that we can construct the likelihood ratio test of this model against the full model.

- For the full model fits, we use the `estimate` and `contrast` statements of `PROC MIXED` to construct the Wald test statistics for different mean slopes, different intercepts, and pairwise comparison of mean slopes for each pair of centers.

PROGRAM:

```
/*******************************************************************

  CHAPTER 9, EXAMPLE 2

  Analysis of the ultrafiltration data by fitting a random
  coefficient model in transmembrane pressure (mmHg)

  -  the repeated measurement factor is transmembrane pressure (tmp)

  -  there is one "treatment" factor, center

  -  the response is ultrafiltration rate (ufr, ml/hr)

  The model for each dialyzer is a straight line.  The intercepts
  and slopes have different means for each center.  The covariance
  matrix D is the same for each center.  The matrix Ri is taken
  to be diagonal with variance sigma^2 for all units.

  We use the RANDOM statement to fit the random coefficient model.

  These data are unbalanced both in the sense that the pressures
  under which each dialyzer is observed are different.

*******************************************************************/

options ls=80 ps=59 nodate; run;

/******************************************************************

  Read in the data set

*******************************************************************/

data ultra; infile 'ultra.dat';
  input subject tmp ufr center;

*  rescale the pressures -- see Chapter 8;

  tmp=tmp/1000;

run;

/*******************************************************************

  Use PROC MIXED to fit the random coefficient model via the
  RANDOM statement.  For all of the fits, we use REML.

  The SOLUTION option in the MODEL statement requests that the
  estimates of the regression parameters be printed.

  In all cases, we take the (2 x 2) matrix D to be unstructured
  (TYPE=UN) in the RANDOM statement.

  The G and GCORR options in the RANDOM statement asks that
  the D matrix and its corresponding correlation matrix
  be printed.  The V and VCORR options ask that the overall
  Sigma matrix be printed (for the first subject or particular
  subjects).

  To fit a random coefficient model, we must specify that both
  intercept and slope are random in the RANDOM statement.

  No REPEATED statement is used because we assume Ri = sigma^2 I,
  which is the default.

*******************************************************************/

*  "Full" model with different intercept, slope for each center;

title 'FULL MODEL, FIT BY REML';
proc mixed data=ultra;
  class center subject;
  model ufr = center center*tmp / noint solution ;
  random intercept tmp / type=un subject=subject g gcorr v vcorr;
  contrast 'diff in slope' center 0 0 0 center*tmp 1 -1 0,
```

```
                          center 0 0 0 center*tmp 1 0 -1 / chisq;
  contrast 'diff in int' center 1 -1 0 center*tmp 0 0 0 ,
                        center 1 0 -1 center*tmp 0 0 0 / chisq;
  estimate 'slope 1 vs 2' center 0 0 0 center*tmp 1 -1 0 ;
  estimate 'slope 1 vs 3' center 0 0 0 center*tmp 1 0 -1 ;
  estimate 'slope 2 vs 3' center 0 0 0 center*tmp 0 1 -1 ;
run;

title 'FULL MODEL, FIT BY ML';
proc mixed method=ml data=ultra;
  class center subject;
  model ufr = center center*tmp / noint solution ;
  random intercept tmp / type=un subject=subject g gcorr v vcorr;
  contrast 'diff in slope' center 0 0 0 center*tmp 1 -1 0,
                          center 0 0 0 center*tmp 1 0 -1 / chisq;
  contrast 'diff in int' center 1 -1 0 center*tmp 0 0 0 ,
                        center 1 0 -1 center*tmp 0 0 0 / chisq;
  estimate 'slope 1 vs 2' center 0 0 0 center*tmp 1 -1 0 ;
  estimate 'slope 1 vs 3' center 0 0 0 center*tmp 1 0 -1 ;
  estimate 'slope 2 vs 3' center 0 0 0 center*tmp 0 1 -1 ;
run;

*  "Reduced" model with different intercepts but same slope for all;
*  centers;

title 'REDUCED MODEL WITH DIFF INTERCEPTS, COMMON SLOPE, FIT BY ML';
proc mixed method=ml data=ultra;
  class center subject;
  model ufr = center tmp / noint solution ;
  random intercept tmp / type=un subject=subject g gcorr v vcorr;
run;
```

OUTPUT: Following the output, we consider the issue of common slopes in several ways.

```
                         FULL MODEL, FIT BY REML                           1

                           The Mixed Procedure

                            Model Information

             Data Set                      WORK.ULTRA
             Dependent Variable            ufr
             Covariance Structure          Unstructured
             Subject Effect                subject
             Estimation Method             REML
             Residual Variance Method      Profile
             Fixed Effects SE Method       Model-Based
             Degrees of Freedom Method     Containment

                         Class Level Information

              Class      Levels    Values

              center          3    1 2 3
              subject        41    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27 28 29 30 31 32 33
                                   34 35 36 37 38 39 40 41

                               Dimensions

                      Covariance Parameters             4
                      Columns in X                      6
                      Columns in Z Per Subject          2
                      Subjects                         41
                      Max Obs Per Subject               5

                         Number of Observations

                  Number of Observations Read             164
                  Number of Observations Used             164
                  Number of Observations Not Used           0

                            Iteration History

          Iteration    Evaluations    -2 Res Log Like       Criterion

                  0              1      1714.69627411
                  1              2      1621.10582541      0.00000580
                  2              1      1621.10190144      0.00000000

                        Convergence criteria met.

                        FULL MODEL, FIT BY REML                           2
```

```
                          The Mixed Procedure

                           Estimated G Matrix

               Row    Effect       subject         Col1         Col2

                 1    Intercept     1           2327.18     -5715.33
                 2    tmp           1          -5715.33        32378

                     Estimated G Correlation Matrix

               Row    Effect       subject         Col1         Col2

                 1    Intercept     1            1.0000      -0.6584
                 2    tmp           1           -0.6584       1.0000

                    Estimated V Matrix for subject 1

               Row        Col1        Col2        Col3        Col4

                 1     2010.79     1271.01     1217.53     1169.94
                 2     1271.01     2255.46     1858.33     2113.31
                 3     1217.53     1858.33     3152.24     3011.76
                 4     1169.94     2113.31     3011.76     4495.01

              Estimated V Correlation Matrix for subject 1

               Row        Col1        Col2        Col3        Col4

                 1      1.0000      0.5968      0.4836      0.3891
                 2      0.5968      1.0000      0.6969      0.6637
                 3      0.4836      0.6969      1.0000      0.8001
                 4      0.3891      0.6637      0.8001      1.0000

                     Covariance Parameter Estimates

                    Cov Parm     Subject    Estimate

                    UN(1,1)      subject     2327.18
                    UN(2,1)      subject    -5715.33
                    UN(2,2)      subject       32378
                    Residual                  683.63

                             Fit Statistics

                 -2 Res Log Likelihood          1621.1
                 AIC (smaller is better)        1629.1
                 AICC (smaller is better)       1629.4
                 BIC (smaller is better)        1636.0

                       FULL MODEL, FIT BY REML                            3

                          The Mixed Procedure

                    Null Model Likelihood Ratio Test

                      DF    Chi-Square      Pr > ChiSq

                       3         93.59          <.0001

                      Solution for Fixed Effects

                                           Standard
 Effect        center    Estimate             Error      DF    t Value    Pr > |t|

 center        1          -174.43           14.9676      82     -11.65      <.0001
 center        2          -172.20           16.9846      82     -10.14      <.0001
 center        3          -151.72           19.2842      82      -7.87      <.0001
 tmp*center    1          4409.53           51.9683      82      84.85      <.0001
 tmp*center    2          4126.00           59.7776      82      69.02      <.0001
 tmp*center    3          4067.73           66.9954      82      60.72      <.0001

                     Type 3 Tests of Fixed Effects

                              Num     Den
               Effect          DF      DF    F Value    Pr > F

               center           3      82     100.17    <.0001
               tmp*center       3      82    5216.74    <.0001

                               Estimates

                                        Standard
     Label            Estimate             Error      DF    t Value    Pr > |t|

     slope 1 vs 2       283.53           79.2090      82       3.58      0.0006
     slope 1 vs 3       341.80           84.7885      82       4.03      0.0001
     slope 2 vs 3      58.2698           89.7872      82       0.65      0.5182
```

```
                                   Contrasts

                  Num     Den
Label              DF      DF    Chi-Square    F Value      Pr > ChiSq    Pr > F

diff in slope       2      82         20.83      10.41          <.0001    <.0001
diff in int         2      82          0.96       0.48          0.6194    0.6211

                            FULL MODEL, FIT BY ML                            4

                              The Mixed Procedure

                               Model Information

             Data Set                     WORK.ULTRA
             Dependent Variable           ufr
             Covariance Structure         Unstructured
             Subject Effect               subject
             Estimation Method            ML
             Residual Variance Method     Profile
             Fixed Effects SE Method      Model-Based
             Degrees of Freedom Method    Containment

                            Class Level Information

              Class      Levels    Values

              center          3    1 2 3
              subject        41    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27 28 29 30 31 32 33
                                   34 35 36 37 38 39 40 41

                                  Dimensions

                      Covariance Parameters             4
                      Columns in X                      6
                      Columns in Z Per Subject          2
                      Subjects                         41
                      Max Obs Per Subject               5

                            Number of Observations

                  Number of Observations Read             164
                  Number of Observations Used             164
                  Number of Observations Not Used           0

                               Iteration History

          Iteration    Evaluations        -2 Log Like       Criterion

                  0             1      1762.75143525
                  1             2      1670.84436023      0.00000724
                  2             1      1670.83930877      0.00000001

                           Convergence criteria met.

                            FULL MODEL, FIT BY ML                            5

                              The Mixed Procedure

                               Estimated G Matrix

               Row    Effect       subject         Col1        Col2

                 1    Intercept     1          2055.33    -5005.31
                 2    tmp           1         -5005.31       29044

                         Estimated G Correlation Matrix

               Row    Effect       subject         Col1        Col2

                 1    Intercept     1           1.0000     -0.6478
                 2    tmp           1          -0.6478      1.0000

                        Estimated V Matrix for subject 1

               Row         Col1         Col2         Col3         Col4

                 1      1880.09      1159.53      1123.70      1091.81
                 2      1159.53      2125.05      1711.25      1950.78
                 3      1123.70      1711.25      2953.75      2768.83
                 4      1091.81      1950.78      2768.83      4179.84

                  Estimated V Correlation Matrix for subject 1

               Row         Col1         Col2         Col3         Col4

                 1       1.0000       0.5801       0.4768       0.3895
                 2       0.5801       1.0000       0.6830       0.6545
```

```
                 3      0.4768      0.6830      1.0000      0.7880
                 4      0.3895      0.6545      0.7880      1.0000

                         Covariance Parameter Estimates

                        Cov Parm     Subject    Estimate

                        UN(1,1)      subject     2055.33
                        UN(2,1)      subject    -5005.31
                        UN(2,2)      subject       29044
                        Residual                  682.93

                                Fit Statistics

                     -2 Log Likelihood              1670.8
                     AIC (smaller is better)        1690.8
                     AICC (smaller is better)       1692.3
                     BIC (smaller is better)        1708.0

                            FULL MODEL, FIT BY ML                             6

                              The Mixed Procedure

                        Null Model Likelihood Ratio Test

                          DF    Chi-Square      Pr > ChiSq

                           3         91.91          <.0001

                          Solution for Fixed Effects

                                      Standard
  Effect        center    Estimate       Error      DF    t Value    Pr > |t|

  center        1          -174.44     14.4204      82     -12.10      <.0001
  center        2          -172.19     16.3531      82     -10.53      <.0001
  center        3          -151.74     18.6268      82      -8.15      <.0001
  tmp*center    1          4409.54     50.0369      82      88.13      <.0001
  tmp*center    2          4125.92     57.5800      82      71.66      <.0001
  tmp*center    3          4067.81     64.6780      82      62.89      <.0001

                         Type 3 Tests of Fixed Effects

                               Num     Den
                Effect          DF      DF    F Value    Pr > F

                center           3      82     107.85    <.0001
                tmp*center       3      82    5618.74    <.0001

                                   Estimates

                                  Standard
      Label           Estimate       Error      DF    t Value    Pr > |t|

      slope 1 vs 2      283.62     76.2833      82       3.72      0.0004
      slope 1 vs 3      341.74     81.7737      82       4.18      <.0001
      slope 2 vs 3     58.1182     86.5950      82       0.67      0.5040

                                   Contrasts

                  Num     Den
Label              DF      DF    Chi-Square    F Value      Pr > ChiSq    Pr > F

diff in slope       2      82         22.43      11.21          <.0001    <.0001
diff in int         2      82          1.03       0.51          0.5986    0.6005

          REDUCED MODEL WITH DIFF INTERCEPTS, COMMON SLOPE, FIT BY ML         7

                              The Mixed Procedure

                               Model Information

             Data Set                     WORK.ULTRA
             Dependent Variable           ufr
             Covariance Structure         Unstructured
             Subject Effect               subject
             Estimation Method            ML
             Residual Variance Method     Profile
             Fixed Effects SE Method      Model-Based
             Degrees of Freedom Method    Containment

                            Class Level Information

              Class      Levels    Values

              center          3    1 2 3
              subject        41    1 2 3 4 5 6 7 8 9 10 11 12 13
                                   14 15 16 17 18 19 20 21 22 23
                                   24 25 26 27 28 29 30 31 32 33
                                   34 35 36 37 38 39 40 41
```

```
                         Dimensions

            Covariance Parameters             4
            Columns in X                      4
            Columns in Z Per Subject          2
            Subjects                         41
            Max Obs Per Subject               5

                   Number of Observations

        Number of Observations Read             164
        Number of Observations Used             164
        Number of Observations Not Used           0

                      Iteration History

 Iteration    Evaluations        -2 Log Like       Criterion

         0              1      1780.28736784
         1              3      1689.51609987      0.00086966
         2              1      1688.81130525      0.00008904
         3              1      1688.74503369      0.00000128
         4              1      1688.74413473      0.00000000

                  Convergence criteria met.

REDUCED MODEL WITH DIFF INTERCEPTS, COMMON SLOPE, FIT BY ML          8

                     The Mixed Procedure

                      Estimated G Matrix

      Row    Effect       subject         Col1         Col2

        1    Intercept     1           3102.51     -9985.70
        2    tmp           1          -9985.70        52598

                Estimated G Correlation Matrix

      Row    Effect       subject         Col1         Col2

        1    Intercept     1            1.0000      -0.7817
        2    tmp           1           -0.7817       1.0000

               Estimated V Matrix for subject 1

      Row        Col1        Col2         Col3         Col4

        1     1938.92     1088.75       931.75       792.02
        2     1088.75     2189.12      1899.08      2250.88
        3      931.75     1899.08      3505.66      3640.26
        4      792.02     2250.88      3640.26      5562.15

         Estimated V Correlation Matrix for subject 1

      Row        Col1        Col2         Col3         Col4

        1      1.0000      0.5285       0.3574       0.2412
        2      0.5285      1.0000       0.6855       0.6451
        3      0.3574      0.6855       1.0000       0.8244
        4      0.2412      0.6451       0.8244       1.0000

                Covariance Parameter Estimates

               Cov Parm     Subject    Estimate

               UN(1,1)      subject     3102.51
               UN(2,1)      subject    -9985.70
               UN(2,2)      subject       52598
               Residual                  685.33

                        Fit Statistics

             -2 Log Likelihood              1688.7
             AIC (smaller is better)        1704.7
             AICC (smaller is better)       1705.7
             BIC (smaller is better)        1718.5

 REDUCED MODEL WITH DIFF INTERCEPTS, COMMON SLOPE, FIT BY ML          9

                     The Mixed Procedure

               Null Model Likelihood Ratio Test

                 DF    Chi-Square      Pr > ChiSq

                  3         91.54          <.0001

                 Solution for Fixed Effects
```

```
                                        Standard
Effect     center     Estimate           Error       DF     t Value     Pr > |t|

center     1           -136.02         12.8851       82      -10.56       <.0001
center     2           -194.43         13.7986       82      -14.09       <.0001
center     3           -187.31         14.8087       82      -12.65       <.0001
tmp                    4230.63         40.4983       40      104.46       <.0001

                        Type 3 Tests of Fixed Effects

                            Num      Den
              Effect         DF       DF     F Value    Pr > F

              center          3       82       90.15    <.0001
              tmp             1       40     10912.8    <.0001
```

INTERPRETATION:

- Comparing to the analysis of these data by ordinary least squares in section 8.8, we see that none of the estimates for $\boldsymbol{\beta}$ in the full model agree with the OLS estimates for the full model. This is not surprising, as these data are **not balanced**.

- In fact, note that the estimates of $\boldsymbol{\beta}$ and their standard errors in the full model in the `Solution for Fixed Effects` table differ slightly for the ML and REML fits. This is to be expected – the "weighting" by the estimated covariance matrices $\widehat{\boldsymbol{\Sigma}}_i$ is slightly different in each case, because the estimates of (the distinct) elements of $\boldsymbol{D}$ and σ^2 are slightly different. This can be seen by inspecting the estimates of $\boldsymbol{D}$ in `Estimated G Matrix` and `Estimated G Correlation Matrix` for each of the ML and REML fits on pages 2 (REML) and page 5 (ML). Similarly, from `Covariance Parameter Estimates` for REML and ML on pages 2 and 5, the estimate of σ^2 may be found (`Residual`). The estimates differ slightly – $\hat{\sigma}^2 = 683.63$ for REML and $\hat{\sigma}^2 = 682.93$ for ML. Note that the estimates of $\boldsymbol{\Sigma}_i$ for the dialyzer $i = 1$ in `Estimated V Matrix for SUBJECT 1` and `Estimated V Correlation Matrix for Subject 1`) are similar for the two fits.

- The results of the `estimate` and `contrast` statements for each fit lead to the same qualitative conclusions. From pages 3 and 6, there is strong evidence according to the Wald (`chisq`) test for difference in slope with 2 degrees of freedom obtained from the `contrast` statement that there is a difference in mean slope for the 3 centers. Here, the $\boldsymbol{L}$ matrix has 2 rows:

$$\boldsymbol{L} = \begin{pmatrix} 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix}.$$

  A `contrast` statement for difference in intercepts, with corresponding $\boldsymbol{L}$ matrix

$$\boldsymbol{L} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \end{pmatrix},$$

  yields in each case a Wald test statistic $T_L = 0.96$ (REML) and 1.03 (ML). Comparing these to a χ^2_2 distribution, it is clear that there is not enough evidence to suggest that the intercepts differ among centers.

The pairwise comparisons of slopes among centers are obtained from the results of the `estimate` statements for each analysis, on pages 3 and 5. Inspection of the results supports the contention that the mean slope for center 1 is different from that for the other two centers. The estimate of this mean slope is 4409.5 (mmHg/100 ml/hr) for each analysis, while those for the other centers are considerably smaller. Thus, it appears that the "typical" rate of change of ultrafiltration rate with transmembrane pressure is faster for dialyzers used at center 1. A possible explanation for this result would be up to the investigators. Perhaps the subject population is different at the first center, or personnel at the first center have different skills operating the devices.

- We may also conduct the test of equal mean slopes via a likelihood ratio test. Here, we use the "full" and "reduced" model results for the fits based on ML. From pages 5 and 8, -2 log-likelihood for the "full" and "reduced" models is 1670.8 and 1688.7, respectively, so that the likelihood ratio test statistic is $1688.7 - 1670.8 = 17.9$. This is to be compared to the χ^2 distribution with $r = 2$ degrees of freedom. As $\chi^2_{2,0.95} = 5.99$, we have strong evidence on the basis of this test to suggest that there is a difference among the mean slopes, which is in agreement with the inference based on the Wald test above.

- For the fit of the "full" model by ML, from page 5, we have $AIC = 1690.8$ and $BIC = 1708.0$. Recall that in section 8.8, we fit the same mean model (although arriving at it from the "population-averaged" perspective) with several different choices of model for $\mathbf{\Sigma}_i$. We may compare those fits to that here, which implies yet another assumption for $\mathbf{\Sigma}_i$, on the basis of AIC and BIC values. The (AIC, BIC) values assuming $\mathbf{\Sigma}_i$ has a compound symmetry and Markov structure, respectively (from pages 4 and 7 of the output in section 8.8), are (1713.5,1727.2) and (1706.0,1719.7), giving support for the "subject-specific" random coefficient modeling approach over the direct, "population-averaged" regression approach in terms of modeling the covariance structure.