Lecture 8: Linear mixed effects models

Adding random effects

A linear model is of the form $Y = X\beta + \mathscr{E}$, where X is a fixed matrix, β is a parameter vector, and $\mathscr{E}$ is an unobserved random error.

In many applications we need to add a random-effect term, which leads to the linear mixed effects model

$$Y = X\beta + Z\alpha + \mathscr{E}$$

where Z is a fixed matrix and α is an unobserved random effect (vector).

The following are two main reasons for adding random effects.

- We want to model the correlation among the errors.
- Random effects present unobserved variables of practical interests.

It is typically assumed that both α and $\mathscr{E}$ have mean 0 and finite covariance matrices, and they are independent; thus,

$$E(Y) = X\beta \qquad \text{and} \qquad \mathrm{Var}(Y) = Z\mathrm{Var}(\alpha)Z' + \mathrm{Var}(\mathscr{E})$$

Example: One-way random effects model

The one-way random effect model

$$Y_{ij} = \mu + A_i + e_{ij}, \qquad j = 1,...,n_i,\ i = 1,...,m,$$

discussed in the last lecture is a special case.

We can derive the MLE's under the one-way random effects model with $A_i \sim N(0, \sigma_a^2)$ and $e_{ij} \sim N(0, \sigma^2)$.

Using the notation previously defined and the form of V^{-1}, we obtain

$$\begin{aligned}
(Y-\mu 1_n)' V^{-1} (Y - \mu 1_n) &= \sum_{i=1}^m (Y_i - \mu 1_{n_i})' (\sigma^2 I_{n_i} + \sigma_a^2 1_{n_i} 1_{n_i}')^{-1} (Y_i - \mu 1_{n_i}) \\
&= \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \mu)^2 - \frac{\sigma_a^2}{\sigma^2} \sum_{i=1}^m \frac{n_i^2 (\bar{Y}_{i\cdot} - \mu)^2}{\sigma^2 + n_i \sigma_a^2} \\
&= \frac{\mathrm{SSE}}{\sigma^2} + \sum_{i=1}^m \frac{n_i (\bar{Y}_{i\cdot} - \mu)^2}{\sigma^2 + n_i \sigma_a^2}, \quad \mathrm{SSE} = \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2
\end{aligned}$$

Since V is block diagonal,

$$|V| = \prod_{i=1}^m |\sigma^2 I_{n_i} + \sigma_a^2 1_{n_i} 1_{n_i}'| = \prod_{i=1}^m \sigma^{2(n_i - 1)} (\sigma^2 + n_i \sigma_a^2)$$

We obtain the likelihood function

$$L(\mu,\sigma^2,\sigma_a^2|Y) = \frac{\exp\left\{-\frac{1}{2}(Y-\mu 1_n)'V^{-1}(Y-\mu 1_n)\right\}}{(2\pi)^{n/2}|V|^{1/2}}$$

$$= \frac{\exp\left\{-\frac{\mathrm{SSE}}{2\sigma^2}-\frac{1}{2}\sum_{i=1}^m \frac{n_i(\bar{Y}_{i\cdot}-\mu)^2}{\sigma^2+n_i\sigma_a^2}\right\}}{(2\pi)^{n/2}\sigma^{n-m}\prod_{i=1}^m(\sigma^2+n_i\sigma_a^2)^{1/2}}$$

Let $\theta=(\mu,\sigma^2,\sigma_a^2)$ and $(\tilde{\mu},\tilde{\sigma}^2,\tilde{\sigma}_a^2)$ be the root of $\partial\log L(\theta|Y)/\partial\theta=0$.

Then we have the following score equations:

$$\tilde{\mu} = \sum_{i=1}^m \frac{n_i\bar{Y}_{i\cdot}}{\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2}\Bigg/\sum_{i=1}^m\frac{n_i}{\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2}$$

$$0 = \frac{\mathrm{SSE}}{\tilde{\sigma}^4}+\sum_{i=1}^m\frac{n_i(\bar{Y}_{i\cdot}-\tilde{\mu})^2}{(\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2)^2}-\frac{n-m}{\tilde{\sigma}^2}-\sum_{i=1}^m\frac{1}{\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2}$$

$$0 = \sum_{i=1}^m\frac{n_i^2(\bar{Y}_{i\cdot}-\tilde{\mu})^2}{(\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2)^2}-\sum_{i=1}^m\frac{n_i}{\tilde{\sigma}^2+n_i\tilde{\sigma}_a^2}$$

There is no explicit formula for this solution.

If $\tilde{\sigma}_a^2 > 0$, then $(\tilde{\mu}, \tilde{\sigma}^2, \tilde{\sigma}_a^2)$ is the MLE of $(\mu, \sigma^2, \sigma_a^2)$.

If $\tilde{\sigma}_a^2 \leq 0$, then the MLE of σ_a^2 is 0, and substituting $\tilde{\sigma}_a^2 = 0$ in the first two score equations we obtain the MLE of μ is $\bar{Y}_{\cdot\cdot}$ and the MLE of σ^2 is

$$\frac{1}{n}\left[\mathrm{SSE} + \sum_{i=1}^{m} n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2\right] = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2$$

In the balance case where $n_i = n_0$ for all i, we have shown that the MLE of μ is $\bar{Y}_{\cdot\cdot}$.

Also, the 3rd score equation becomes

$$\tilde{\sigma}^2 + n_0\tilde{\sigma}_a^2 = \frac{n_0}{m}\sum_{i=1}^{m}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

and because of the 3rd equation, the 2nd score equation becomes

$$\tilde{\sigma}^2 = \frac{\mathrm{SSE}}{n-m}$$

Then

$$\tilde{\sigma}_a^2 = \frac{1}{m}\sum_{i=1}^{m}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 - \frac{\mathrm{SSE}}{n_0(n-m)}$$

The MLE of $(\mu, \sigma^2, \sigma_a^2)$ in the balanced case is $(\bar{Y}_{\cdot\cdot}, \tilde{\sigma}^2, \max(\tilde{\sigma}_a^2, 0))$.

The restricted maximum likelihood (REML) estimators

To introduce the idea, let's first consider the example of $X = (X_1, ..., X_n)$, where X_i's are iid $N(\mu, \sigma^2)$ so that the likelihood is

$$\begin{aligned} L(\mu, \sigma^2|X) &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{(\bar{X}-\mu)^2}{2\sigma^2/n} - \frac{(n-1)S^2}{2\sigma^2}\right\} \\ &= L_1(\mu|\bar{X}, \sigma^2)L_2(\sigma^2|S^2) \end{aligned}$$

where

$$L_1(\mu|\bar{X}, \sigma^2) = \frac{C_1}{\sigma} \exp\left\{-\frac{(\bar{X}-\mu)^2}{2\sigma^2/n}\right\}$$

$$L_2(\sigma^2|S^2) = \frac{C_2}{\sigma^{(n-1)}} \exp\left\{-\frac{(n-1)S^2}{2\sigma^2}\right\}$$

and C_1 and C_2 are constants not depending on parameters.

Maximizing L_1 over μ we obtain the maximum $\widehat{\mu} = \bar{X}$ and maximizing L_2 we obtain the maximum $\widehat{\sigma}^2 = S^2$.

The REML estimator of (μ, σ^2) is $(\bar{X}, S^2)$.

The MLE of (μ, σ^2) is $(\bar{X}, (n-1)S^2/n)$.

REML estimators in balanced one-way random effects model

From the previous derivation we know that the likelihood under the balanced ($n_i = n_0$ for all i) one-way random effects model is

$$L(\mu, \sigma^2, \sigma_a^2 | Y) = \frac{\exp\left\{ -\frac{\mathrm{SSE}}{2\sigma^2} - \frac{\mathrm{SSA}}{2(\sigma^2 + n_0\sigma_a^2)} - \frac{n(\bar{Y}_{\cdot\cdot} - \mu)^2}{2(\sigma^2 + n_0\sigma_a^2)} \right\}}{(2\pi)^{n/2}\sigma^{n-m}(\sigma^2 + n_0\sigma_a^2)^{m/2}}$$

where

$$\mathrm{SSA} = \sum_{i=1}^m n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

Maximizing

$$L_1(\mu | \bar{Y}_{\cdot\cdot}, \sigma^2, \sigma_a^2) = \frac{1}{(\sigma^2 + n_0\sigma_a^2)^{1/2}} \exp\left\{ -\frac{n(\bar{Y}_{\cdot\cdot} - \mu)^2}{2(\sigma^2 + n_0\sigma_a^2)} \right\}$$

we obtain that the REML estimator of μ is $\widehat{\mu} = \bar{Y}_{\cdot\cdot}$; maximizing

$$L_2(\sigma^2, \sigma_a^2 | \mathrm{SSE}, \mathrm{SSA}) = \frac{\exp\left\{ -\frac{\mathrm{SSE}}{2\sigma^2} - \frac{\mathrm{SSA}}{2(\sigma^2 + n_0\sigma_a^2)} \right\}}{\sigma^{n-m}(\sigma^2 + n_0\sigma_a^2)^{(m-1)/2}}$$

we obtain that the REML estimators of σ^2 and σ_a^2 are, respectively,

$$\widehat{\sigma}^2 = \frac{\mathrm{SSE}}{n-m} \quad \text{and} \quad \check{\sigma}_a^2 = \max(\tilde{\sigma}_a^2,\ 0), \quad \check{\sigma}_a^2 = \frac{\mathrm{SSA}}{n - n_0} - \frac{\mathrm{SSE}}{n_0(n-m)}$$

In fact, $\check{\sigma}_a^2$ is the popular ANOVA estimator, although it may take negative values.

Comparing the MLE derived previously with the REML estimators, we find that the MLE and REML estimators of μ and σ^2 are the same, but the estimators of σ_a^2 are the positive parts of two different estimators,

$$\tilde{\sigma}_a^2 = \frac{\mathrm{SSA}}{n} - \frac{\mathrm{SSE}}{n_0(n-m)} \quad \text{and} \quad \check{\sigma}_a^2 = \frac{\mathrm{SSA}}{n-n_0} - \frac{\mathrm{SSE}}{n_0(n-m)}$$

The difference is in the denominator of the first term.

Consider the expectations of these two estimators.

First, since

$$\mathrm{SSE} = \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 = \sum_{i=1}^m (n_i - 1) S_i^2$$

where S_i is the sample variance based on $Y_{i1}, ..., Y_{in_i}$,

$$E(\hat{\sigma}^2) = E\left(\frac{\mathrm{SSE}}{n-m}\right) = \frac{1}{n-m}\sum_{i=1}^m (n_i-1)E(S_i^2) = \frac{\sigma^2}{n-m}\sum_{i=1}^m (n_i-1) = \sigma^2$$

Hence, the MLE or REML estimator of σ^2 is unbiased (UMVUE).

Note that

$$\bar{Y}_{i\cdot} = \mu + A_i + \frac{1}{n_i}\sum_{j=1}^{n_i} \varepsilon_{ij} \sim N(\mu, \sigma_a^2 + n_i^{-1}\sigma^2)$$

When $n_i = n_0$ for all i, $\bar{Y}_{1\cdot}, ..., \bar{Y}_{m\cdot}$ are iid from $N(\mu, \sigma_a^2 + n_0^{-1}\sigma^2)$ and, hence

$$\frac{\text{SSA}}{n_0(m-1)} = \frac{1}{m-1}\sum_{i=1}^{m} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

is the sample variance of $\bar{Y}_{1\cdot}, ..., \bar{Y}_{m\cdot}$ and

$$E\left(\frac{\text{SSA}}{n-n_0}\right) = E\left(\frac{\text{SSA}}{n_0(m-1)}\right) = \sigma_a^2 + \frac{\sigma^2}{n_0}$$

This implies

$$E(\check{\sigma}_a^2) = E\left(\frac{\text{SSA}}{n-n_0}\right) - E\left(\frac{\text{SSE}}{n_0(n-m)}\right) = \sigma_a^2 + \frac{\sigma^2}{n_0} - \frac{\sigma^2}{n_0} = \sigma_a^2$$

i.e., the ANOVA estimator $\check{\sigma}_a^2$ is unbiased (UMVUE).

On the other hand,

$$E(\tilde{\sigma}_a^2) = E\left(\frac{\text{SSA}}{n} - \frac{\text{SSE}}{n_0(n-m)}\right) = \frac{n-n_0}{n}\left(\sigma_a^2 + \frac{\sigma^2}{n_0}\right) - \frac{\sigma^2}{n_0}$$

Thus, $\tilde{\sigma}_a^2$ is biased and the bias is

$$E(\tilde{\sigma}_a^2) - \sigma_a^2 = -\frac{n_0\sigma_a^2}{n} - \frac{\sigma^2}{n} = -\frac{\sigma_a^2}{m} - \frac{\sigma^2}{n}$$

This bias becomes a serious issue when m is not large.

REML estimators in unbalanced one-way random effects model

For the unbalanced case, the previously derived likelihood is

$$L(\mu, \sigma^2, \sigma_a^2 | Y) = \frac{\exp\left\{ -\frac{\mathrm{SSE}}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^{m} \frac{n_i(\bar{Y}_{i\cdot} - \mu)^2}{\sigma^2 + n_i\sigma_a^2} \right\}}{(2\pi)^{n/2}\sigma^{n-m}\prod_{i=1}^{m}(\sigma^2 + n_i\sigma_a^2)^{1/2}}$$

$$= \frac{\exp\left\{ -\frac{\mathrm{SSE}}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^{m} \frac{n_i(\bar{Y}_{i\cdot} - \tilde{\mu})^2}{\sigma^2 + n_i\sigma_a^2} - \frac{1}{2}\sum_{i=1}^{m} \frac{n_i(\tilde{\mu} - \mu)^2}{\sigma^2 + n_i\sigma_a^2} \right\}}{(2\pi)^{n/2}\sigma^{n-m}\prod_{i=1}^{m}(\sigma^2 + n_i\sigma_a^2)^{1/2}}$$

where

$$\tilde{\mu} = \sum_{i=1}^{m} \frac{n_i\bar{Y}_{i\cdot}}{\sigma^2 + n_i\sigma_a^2} \Bigg/ \sum_{i=1}^{m} \frac{n_i}{\sigma^2 + n_i\sigma_a^2}$$

For each fixed (σ^2, σ_a^2),

$$\tilde{\mu} \sim N\left(\mu, \tau^2\right), \quad \tau^2 = \left(\sum_{i=1}^{m} \frac{n_i}{\sigma^2 + n_i\sigma_a^2}\right)^{-1}$$

and $\tilde{\mu}$ maximizes

$$L_1(\mu|Y, \sigma^2, \sigma_a^2) = \frac{1}{(2\pi)^{1/2}\tau} \exp\left\{-\frac{1}{2}\sum_{i=1}^{m} \frac{n_i(\tilde{\mu}-\mu)^2}{\sigma^2 + n_i\sigma_a^2}\right\}$$

Let

$$L_2(\sigma^2, \sigma_a^2|\mathrm{SSE}, \bar{Y}_{i\cdot} - \tilde{\mu}, i=1,...,m) = \frac{\exp\left\{-\frac{\mathrm{SSE}}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^{m} \frac{n_i(\bar{Y}_{i\cdot}-\tilde{\mu})^2}{\sigma^2+n_i\sigma_a^2}\right\}}{(2\pi)^{(n-1)/2}\sigma^{n-m}\prod_{i=1}^{m}(\sigma^2 + n_i\sigma_a^2)^{1/2}\tau^{-1}}$$

Then

$$L(\mu, \sigma^2, \sigma_a^2|Y) = L_1(\mu|\tilde{\mu}, \sigma^2, \sigma_a^2)L_2(\sigma^2, \sigma_a^2|\mathrm{SSE}, \bar{Y}_{i\cdot} - \tilde{\mu}, i=1,...,m)$$

Note that the second factor in the product depends on $\tilde{\mu}$, due to unbalancedness.

Like the MLE, the REML estimators have to be calculated iteratively.

At each iteration, first treat the previously calculated estimates of σ^2 and σ_a^2 as known and calculate $\tilde{\mu}$ (it is a WLSE).

Then treat $\tilde{\mu}$ as known and calculate MLE of (σ^2, σ_a^2) by maximizing

$$L_2(\sigma^2, \sigma_a^2 | \mathrm{SSE}, \bar{Y}_{i\cdot} - \tilde{\mu}, i = 1, ..., m)$$

This step needs iteration, since no explicit solution exists due to unbalancedness.

Also, we need to restrict the estimate of σ_a^2 to $[0, \infty)$.

The transformation approach to derive REML estimators

The matrix form of the one way balanced random effects model is

$$Y = \mu 1_n + Z\alpha + \mathscr{E}$$

where

$$Z = \begin{pmatrix} 1_{n_0} & 0 & \cdots & 0 \\ 0 & 1_{n_0} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1_{n_0} \end{pmatrix} \qquad \alpha = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}$$

Note that $H_n = I_n - n^{-1}1_n1_n'$ is an $n \times n$ projection matrix of rank $n-1$.

Let $\tilde{H}_n$ be the $(n-1)\times n$ matrix obtained by deleting the first row of H_n and

$$C = \begin{pmatrix} 1_n' \\ \tilde{H}_n \end{pmatrix}$$

Because $\tilde{H}_n$ is of rank $n-1$ and

$$H_n 1_n = (I_n - n^{-1}1_n1_n')1_n = 1_n - 1_n = 0$$

we know that $\tilde{H}_n 1_n = 0$ so that C is of rank n and observing Y is equivalent to observing CY.

The first component of CY is

$$\begin{aligned} \bar{Y}_{\cdot\cdot} &= n^{-1}1_n'Y \\ &= \mu + n^{-1}1_n'Z\alpha + n^{-1}1_n'\mathscr{E} \\ &= \mu + \frac{1}{n}\sum_{i=1}^m n_0 A_i + \frac{1}{n}\sum_{i=1}^m\sum_{j=1}^{n_0}\varepsilon_{ij} \\ &\sim N\left(\mu,\ \frac{\sigma^2 + n_0\sigma_a^2}{n}\right) \end{aligned}$$

Because $\tilde{H}_n 1_n = 0$, the last $n-1$ component of CY is

$$\tilde{H}_n Y = \tilde{H}_n Z\alpha + \tilde{H}_n \mathscr{E} \sim N(0,\ \tilde{H}_n Z \mathrm{Var}(\alpha) Z' \tilde{H}_n' + \sigma^2 \tilde{H}_n \tilde{H}_n')$$

which does not involve μ, and $\bar{Y}_{\cdot\cdot}$ and $\tilde{H}_n Y$ are independent.

Hence, the entire likelihood can have the decomposition

$$L(\mu, \sigma^2, \sigma_a^2 | Y) = L_1(\mu | \bar{Y}_{\cdot\cdot}, \sigma^2, \sigma_a^2) L_2(\sigma^2, \sigma_a^2 | \tilde{H}_n Y)$$

where the second factor actually depends SSE and SSA only.

MLE in a general linear mixed effects model

Consider a general linear mixed effects model

$$Y = X\beta + Z\alpha + \mathscr{E}$$

as defined in the beginning.

We assume that the matrix X is of rank $=$ the dimension of β and denote $V_\alpha = \mathrm{Var}(\alpha)$ and $V = \mathrm{Var}(Y) = ZV_\alpha Z' + \sigma^2 I_n$.

Assuming that α and $\mathscr{E}$ are independently normal, and letting

$$\widehat{\beta}_{V^{-1}} = (X'V^{-1}X)^{-1} X'V^{-1} Y$$

we obtain the likelihood as

$$L(\beta,\sigma^2,V_\alpha|Y) = \frac{\exp\left\{-\frac{(Y-X\beta)'V^{-1}(Y-X\beta)}{2}\right\}}{(2\pi)^{n/2}|V|^{1/2}}$$

$$= \frac{\exp\left\{-\frac{(Y-X\widehat{\beta}_{V^{-1}})'V^{-1}(Y-X\widehat{\beta}_{V^{-1}})}{2} - \frac{(\beta-\widehat{\beta}_{V^{-1}})'X'V^{-1}X(\beta-\widehat{\beta}_{V^{-1}})}{2}\right\}}{(2\pi)^{n/2}|V|^{1/2}}$$

because

$$(Y-X\beta)'V^{-1}(Y-X\beta) = (Y-X\widehat{\beta}_{V^{-1}})'V^{-1}(Y-X\widehat{\beta}_{V^{-1}}) + (\beta-\widehat{\beta}_{V^{-1}})'X'V^{-1}X(\beta-\widehat{\beta}_{V^{-1}})$$

since, by the definition of $\widehat{\beta}_{V^{-1}}$,

$$(Y-X\widehat{\beta}_{V^{-1}})'V^{-1}X(\beta-\widehat{\beta}_{V^{-1}}) = Y'V^{-1}X - \widehat{\beta}'_{V^{-1}}X'V^{-1}X(\beta-\widehat{\beta}_{V^{-1}}) = 0$$

Hence, the MLE of β and V should be $\widehat{\beta}_{\mathrm{ML}} = \widehat{\beta}_{\widehat{V}_{\mathrm{ML}}^{-1}}$ and $\widehat{V}_{\mathrm{ML}}$, where $\widehat{V}_{\mathrm{ML}}$ is the solution to

$$\max_V \frac{1}{|V|^{1/2}} \exp\left\{-\frac{(Y-X\widehat{\beta}_{V^{-1}})'V^{-1}(Y-X\widehat{\beta}_{V^{-1}})}{2}\right\}$$

REML estimators in a general linear mixed effects model

The REML estimator of V is the solution to

$$\max_V \frac{1}{|V|^{1/2}|X'V^{-1}X|^{1/2}} \exp\left\{-\frac{(Y - X\widehat{\beta}_{V^{-1}})'V^{-1}(Y - X\widehat{\beta}_{V^{-1}})}{2}\right\}$$

In view of

$$\widehat{\beta}_{V^{-1}} \sim N\left(\beta,\ (X'V^{-1}X)^{-1}\right)$$

the REML estimator is the MLE with adjustment factor $|X'V^{-1}X|^{1/2}$, which accounts the estimation of β when we estimate V.

In general, both ML and REML estimators do not have explicit forms, and we can iterate between the estimation of β and V.

We usually need some information about $V_\alpha = \mathrm{Var}(\alpha)$ so that the function to be maximized can be simplified.

The one way random effects model is an example.

As we discussed in the last lecture, if $X(X'X)^{-1}X'V$ is symmetric, which is equivalent to $X(X'X)^{-1}X'V_\alpha$ is symmetric, then $\widehat{\beta}_{V^{-1}} = (X'X)^{-1}X'Y = \widehat{\beta}$ is the LSE and has an explicit form not depending on estimators of V.