Conditional Distribution of the Random Effects

$$\begin{aligned} \text{Model: } & \mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{A}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i, \\ & \boldsymbol{\epsilon}_i = \{\epsilon_{i,1}, \ldots, \epsilon_{i,n_i}\}, \quad \epsilon_{i,j} \sim \text{Normal}(0, \sigma^2), \\ \text{Random effect: } & \boldsymbol{\eta}_i \sim \text{Normal}(0, \boldsymbol{\Omega}) \end{aligned}$$

and n_i is the length of $\mathbf{y}_i$.

Step 1: Joint Distribution of $(\mathbf{y}_i, \boldsymbol{\eta}_i)$

We have $E(\mathbf{y}_i) = E(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{A}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i) = \mathbf{X}_i\boldsymbol{\beta}$ and $V(\mathbf{y}_i) = E\{(\mathbf{y}_i - E(\mathbf{y}_i))(\mathbf{y}_i - E(\mathbf{y}_i))^T\} = E\{(\mathbf{A}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i)(\mathbf{A}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i)^T\} = E(\mathbf{A}_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i^T\mathbf{A}_i^T) + E(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i^T) = \mathbf{A}_i\boldsymbol{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i}$ since $\boldsymbol{\eta}_i$ and $\boldsymbol{\epsilon}_i$ are independent by assumption and thus, $\text{Cov}(\boldsymbol{\eta}_i, \boldsymbol{\epsilon}_i) = 0$.

Then, $\text{Cov}(\mathbf{y}_i, \boldsymbol{\eta}_i) = E(\mathbf{A}_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i^T) = \mathbf{A}_i\boldsymbol{\Omega}$.

In parallel, $E(\boldsymbol{\eta}_i) = 0$ and $\text{V}(\boldsymbol{\eta}_i) = \boldsymbol{\Omega}$, by assumption.

The joint distribution of $\mathbf{y}_i$ and $\boldsymbol{\eta}_i$ is then:

$$\begin{pmatrix} \mathbf{y}_i \\ \boldsymbol{\eta}_i \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mathbf{X}_i\boldsymbol{\beta} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{A}_i\boldsymbol{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i} & \mathbf{A}_i\boldsymbol{\Omega} \\ (\mathbf{A}_i\boldsymbol{\Omega})^T & \boldsymbol{\Omega} \end{pmatrix} \right).$$

Step 2: Conditional Distribution of $\boldsymbol{\eta}_i \mid \mathbf{y}_i$

Using the formulas from `https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Conditional_distributions`, we have that the conditional distribution of $\boldsymbol{\eta}_i$ given $\mathbf{y}_i$ is:

$$\boldsymbol{\eta}_i \mid \mathbf{y}_i \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\eta}_i|\mathbf{y}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}_i|\mathbf{y}_i}),$$

where

$$\boldsymbol{\mu}_{\boldsymbol{\eta}_i|\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}} = (\mathbf{A}_i\boldsymbol{\Omega})^T(\mathbf{A}_i\boldsymbol{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta}),$$

and

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}_i|\mathbf{y}_i} = \boldsymbol{\Omega} - (\mathbf{A}_i\boldsymbol{\Omega})^T(\mathbf{A}_i\boldsymbol{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}(\mathbf{A}_i\boldsymbol{\Omega}).$$

Step 3: Simplifying $\mathbf{\Sigma}_{\boldsymbol{\eta}_i|\mathbf{y}_i}$

We define:

$$\mathbf{Q}_i = \mathbf{\Omega}^{-1} + \mathbf{A}_i\mathbf{A}_i^T/\sigma^2.$$

Then, using the Sherman-Morrison-Woodbury identity (`https://en.wikipedia.org/wiki/Sherman%E2%80%93Morrison_formula#Generalization_(Woodbury_matrix_identity)`), we can express:

$$\mathbf{\Sigma}_{\boldsymbol{\eta}_i|\mathbf{y}_i} = \mathbf{Q}_i^{-1}.$$

Then, see that

$$\begin{aligned}
\mathbf{Q}_i^{-1}\mathbf{A}_i^T/\sigma^2 &= \{\mathbf{\Omega} - (\mathbf{A}_i\mathbf{\Omega})^T(\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}(\mathbf{A}_i\mathbf{\Omega})\}\mathbf{A}_i^T/\sigma^2 \\
&= \mathbf{\Omega}\mathbf{A}_i^T/\sigma^2 - \mathbf{\Omega}\mathbf{A}_i^T(\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}(\mathbf{A}_i\mathbf{\Omega})\mathbf{A}_i^T/\sigma^2 \\
&= \mathbf{\Omega}\mathbf{A}_i^T/\sigma^2\{\sigma^2\mathbf{I}_{n_i} - (\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T\} \\
&= \mathbf{\Omega}\mathbf{A}_i^T/\sigma^2(\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}\{(\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i}) - \mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T\} \\
&= (\mathbf{A}_i\mathbf{\Omega})^T(\mathbf{A}_i\mathbf{\Omega}\mathbf{A}_i^T + \sigma^2\mathbf{I}_{n_i})^{-1}
\end{aligned}$$

Thus, the conditional distribution becomes:

$$\boldsymbol{\eta}_i \mid \mathbf{y}_i \sim \mathcal{N}\left(\mathbf{Q}_i^{-1}\mathbf{A}_i^T(\mathbf{y}_i - \mathbf{X}\boldsymbol{\beta})/\sigma^2, \mathbf{Q}_i^{-1}\right).$$