Probability Model For A Multivariate Hierarchical Linear Regression Model

Nan Wu
nanw@udel.edu

We consider a two level hierarchical linear regression model in which the data y are organized into J distinct groups. Each individual observation y_n has a predictor row vector x_n of size K with $x_{n,1} = 1$ to unify the notation. Each group j gets its own coefficient vector β_j of size K, which is predictable from its group-level predictors u_j of size L.

1 Multivariate Hierarchical Linear Regression Model

The two level hierarchical linear regression model is described by the following two level regression equations.

$$y_n = \sum_{k=1}^{K} \beta_{j_n k} x_{nk} + e_n \tag{1}$$

$$\beta_j = \sum_{l=1}^{L} \gamma_l u_{jl} + \epsilon_j \tag{2}$$

where $x \in \mathbb{R}^{N \times K}$ is the individual-level predictor matrix; $u \in \mathbb{R}^{J \times L}$ is the group-level predictor matrix; $\beta_j \in \mathbb{R}^{1 \times K}$ is the individual-level coefficient vector by group; $\gamma_l \in \mathbb{R}^{1 \times K}$ is the group-level coefficient vector; $\epsilon_j \in \mathbb{R}^{1 \times K}$ is the group-level residual vector; $j_n \in \{1, \ldots, J\}$; $n = 1, \ldots, N$; and $j = 1, \ldots, J$.

2 Probability Model

Assume $e_n \sim \text{Normal}(0, \sigma)$, and $\epsilon_j \sim \text{MultiNormal}(0, \Sigma)$, then we have

$$y_n \sim \text{Normal}\left(\sum_{k=1}^{K} \beta_{j_n k} x_{nk}, \sigma\right) \tag{3}$$

$$\beta_j \sim \text{MultiNormal}\left(\sum_{l=1}^{L} \gamma_l u_{jl}, \Sigma\right) \tag{4}$$

Then we must give priors to σ, γ, and Σ. We may assume an improper prior for σ, and the following independent weakly informative prior for γ.

$$\gamma_{lk} \sim \text{Normal}\left(0, 5\right)$$

where $l = 1, \ldots, L$, and $k = 1, \ldots, K$.

We decompose Σ into a scale and a correlation matrix as below

$$\Sigma = \text{diag\_matrix}\left(\tau\right)\Omega\text{diag\_matrix}\left(\tau\right) \tag{5}$$

where $\Sigma \in \mathbb{R}^{K \times K}$ is a correlation matrix and $\tau \in \mathbb{R}^{K}$ is the vector of coefficient scales that

$$\begin{aligned} \tau_k &= \sqrt{\Sigma_{kk}} \\ \Omega_{ij} &= \frac{\Sigma_{ij}}{\tau_i \tau_j} \end{aligned}$$

We may assume the following priors for τ and Ω:

$$\begin{aligned} \tau_k &\sim \text{Cauchy}_{+}\left(0, 2.5\right) \\ \Omega &\sim \text{LKJCorr}\left(2\right) \end{aligned}$$

where $k = 1, \ldots, K$.

Putting it all together, we have the probability model for this two level hierarchical linear regression model as:

$$\begin{aligned} y_n &\sim \text{Normal}\left(\sum_{k=1}^{K} \beta_{j_n k} x_{nk}, \sigma\right) \\ \sigma &\sim \text{Uniform}\left(0, \infty\right) \\ \beta_j &\sim \text{MultiNormal}\left(\sum_{l=1}^{L} \gamma_l u_{jl}, \Sigma\right) \\ \gamma_{lk} &\sim \text{Normal}\left(0, 5\right) \\ \Sigma &= \text{diag\_matrix}\left(\tau\right)\Omega\text{diag\_matrix}\left(\tau\right) \\ \tau_k &\sim \text{Cauchy}_{+}\left(0, 2.5\right) \\ \Omega &\sim \text{LKJCorr}\left(2\right) \end{aligned}$$

where $j_n \in \{1, \ldots, J\}$.