Distributions
2020-07-15

Our model:

$$\boldsymbol{y_{ki}} = \boldsymbol{X_i}(\boldsymbol{\beta_k} + \boldsymbol{b_{ki}} + \boldsymbol{\Gamma_k}(\boldsymbol{\alpha' x_i})) + \boldsymbol{\epsilon_{ki}} \tag{1}$$

where

- $\boldsymbol{X}_i$ is the $n_t \times t$ dimension design matrix for subject i;
- $\boldsymbol{x}_i = \begin{pmatrix} x_{i1} & \dots & x_{ip} \end{pmatrix}'$ is the $p \times 1$ vector of predictors for subject i; x_{ij} presents the jth predictor for the ith subject; $\boldsymbol{x}_i \sim MVN(\boldsymbol{0}, \boldsymbol{\Sigma}_x)$
- $\boldsymbol{\beta}_k$ is the fixed effect coefficient with dimension $t \times 1$;
- $\boldsymbol{\Gamma}_{kj}$ is the fixed effect coefficient with dimension $t \times 1$;
- $\boldsymbol{\alpha}$ is a $p \times 1$ vector.
- $\boldsymbol{b}_{ki} \sim N(0, \boldsymbol{D}_k)$, $\boldsymbol{D}_k$ is a $t \times t$ matrix, with $\frac{t(t+1)}{2}$ parameters.
- $\boldsymbol{\epsilon}_{ki} \sim N(0, \sigma_k^2)$

Therefore, the expectation and variance of outcome Y_k given the predictors is

$$E(\boldsymbol{Y}_k|\boldsymbol{x}) = \boldsymbol{X\beta}_k + \boldsymbol{X\Gamma_k\alpha' x} \tag{2}$$

$$Var(\boldsymbol{Y}_k|\boldsymbol{x}) = \boldsymbol{XD_kX'} + \sigma_k^2\boldsymbol{I} \tag{3}$$

Therefore, the conditional distribution of outcome $\boldsymbol{Y}_k|\boldsymbol{x}$ is

$$\boldsymbol{Y_k}|\boldsymbol{x} \sim MVN(\boldsymbol{X\beta}_k + \boldsymbol{X\Gamma_k\alpha' x}, \boldsymbol{XD_kX'} + \sigma_k^2\boldsymbol{I}) \tag{4}$$

Besides, the expectation of variance of the outcome is

$$E(\boldsymbol{Y}_k) = E(E(\boldsymbol{Y}_k|\boldsymbol{x})) = \boldsymbol{X\beta}_k \tag{5}$$

since $E(\boldsymbol{x}) = \boldsymbol{0}$.

$$\begin{aligned} Var(Y_k) =& E(Var(\boldsymbol{Y}_k|\boldsymbol{x})) + Var(E(\boldsymbol{Y}_k|\boldsymbol{x})) \\ =& E(\boldsymbol{XD_kX'} + \sigma_k^2\boldsymbol{I}) + Var(\boldsymbol{X\beta}_k + \boldsymbol{X\Gamma_k\alpha' x}) \\ =& \boldsymbol{XD_kX'} + \sigma_k^2\boldsymbol{I} + \boldsymbol{X\Gamma_k\alpha'}\boldsymbol{\Sigma}_x(\boldsymbol{X\Gamma_k\alpha'})' \\ =& \boldsymbol{X}(\boldsymbol{D_k} + \boldsymbol{\Gamma_k\alpha'\Sigma_x\alpha\Gamma_k'})\boldsymbol{X'} + \sigma_k^2\boldsymbol{I} \end{aligned} \tag{6}$$

Therefore, the distribution of $\boldsymbol{Y}_k$ is

$$\boldsymbol{Y}_k \sim MVN(\boldsymbol{X\beta}_k, \boldsymbol{X}(\boldsymbol{D_k} + \boldsymbol{\Gamma_k\alpha'\Sigma_x\alpha\Gamma_k'})\boldsymbol{X'} + \sigma_k^2\boldsymbol{I}) \tag{7}$$

The joint distributions between $\boldsymbol{Y}_k$ and $\boldsymbol{x}$ is

$$\begin{aligned} f(\boldsymbol{y}_k, \boldsymbol{x}) =& f(\boldsymbol{Y}_k|\boldsymbol{x})f(\boldsymbol{x}) \\ =& \frac{1}{\sqrt{(2\pi)^{n_t}|\boldsymbol{\Sigma}_k|}}\exp(-\frac{1}{2}(\boldsymbol{y}_k - \boldsymbol{\mu}_k)'\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{y}_k - \boldsymbol{\mu}_k)) \times \frac{1}{\sqrt{(2\pi)^p|\boldsymbol{\Sigma}_x|}}\exp(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{\Sigma}_x^{-1}\boldsymbol{x}) \end{aligned} \tag{8}$$

where $\boldsymbol{\Sigma}_k = \boldsymbol{XD_kX'} + \sigma_k^2\boldsymbol{I}$, $\boldsymbol{\mu}_k = \boldsymbol{X}(\boldsymbol{\beta_k} + \boldsymbol{\Gamma_k}(\boldsymbol{\alpha' x}))$

The covariance matrix between $\boldsymbol{Y}_k$ and $\boldsymbol{x}$ is

$$\begin{aligned} Cov(\boldsymbol{Y}_k, \boldsymbol{x}) =& Cov(\boldsymbol{X}(\boldsymbol{\beta_k} + \boldsymbol{b_k} + \boldsymbol{\Gamma_k}(\boldsymbol{\alpha' x})) + \boldsymbol{\epsilon_k}, \boldsymbol{x}) \\ =& Cov(\boldsymbol{X\Gamma_k\alpha' x}, \boldsymbol{x}) \\ =& \boldsymbol{X\Gamma_k\alpha'}Cov(\boldsymbol{x}, \boldsymbol{x}) \\ =& \boldsymbol{X\Gamma_k\alpha'}\boldsymbol{\Sigma}_x \quad (n \times p) \\ Cov(\boldsymbol{x}, \boldsymbol{Y}_k) =& \boldsymbol{\Sigma_x\alpha\Gamma_k'X'} \end{aligned} \tag{9}$$

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y}_k \end{pmatrix} \sim MVN(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{X\beta_k} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_x & \boldsymbol{\Sigma_x\alpha\Gamma_k'X'} \\ \boldsymbol{X\Gamma_k\alpha'}\boldsymbol{\Sigma}_x & \boldsymbol{X}(\boldsymbol{D_k} + \boldsymbol{\Gamma_k\alpha'\Sigma_x\alpha\Gamma_k'})\boldsymbol{X'} + \sigma_k^2\boldsymbol{I} \end{pmatrix}) \tag{10}$$

Likelihood optimization

1. Optimize the conditional log-likelihood function $\log(f(\boldsymbol{y}|\boldsymbol{x}))$

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{x}) =& - n \cdot n_t \log(2\pi) - \frac{n}{2}\log(|\boldsymbol{\Sigma}_1|) - \frac{n}{2}\log(|\boldsymbol{\Sigma}_2|) \\ & - \sum_{i=1}^{n} \frac{1}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}_{1i})'\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_{1i}) \\ & - \sum_{i=n+1}^{2n} \frac{1}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}_{2i})'\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_{2i}) \end{aligned} \tag{11}$$

where $\boldsymbol{\Sigma}_k = \boldsymbol{X}\boldsymbol{D_k}\boldsymbol{X}' + \sigma_k^2\boldsymbol{I}$, $\boldsymbol{\mu}_k = \boldsymbol{X(\beta_k + \Gamma_k(\alpha' x))}$

2. Optimize the log-likelihood function $\log(f(\boldsymbol{y}))$.

$$\begin{aligned} l(\boldsymbol{\theta}) =& - n \cdot n_t \log(2\pi) - \frac{n}{2}\log(|\boldsymbol{\Sigma}_{y1}|) - \frac{n}{2}\log(|\boldsymbol{\Sigma}_{y2}|) \\ & - \sum_{i=1}^{n} \frac{1}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}_{y1i})'\boldsymbol{\Sigma}_{y1}^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_{y1i}) \\ & - \sum_{i=n+1}^{2n} \frac{1}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}_{y2i})'\boldsymbol{\Sigma}_{y2}^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_{y2i}) \end{aligned} \tag{12}$$

where $\boldsymbol{\mu}_{yki} = \boldsymbol{X}_i\boldsymbol{\beta}_k$, $\boldsymbol{\Sigma}_{yk} = \boldsymbol{X}_i(\boldsymbol{D_k + \Gamma_k\alpha'\Sigma_x\alpha\Gamma_k'})\boldsymbol{X}_i' + \sigma_k^2\boldsymbol{I}$

3. Optimize the mean value of expectation of conditional log-likelihood function.

$$\begin{aligned} E(l(\boldsymbol{\theta})) =& E(E(l(\boldsymbol{\theta}|\boldsymbol{x}))) \propto -\frac{n}{2}\log(|\boldsymbol{\Sigma}_1|) - \frac{n}{2}\log(|\boldsymbol{\Sigma}_2|) - \frac{n}{4}tr(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1) - \frac{n}{4}tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2) \\ & - \frac{n}{4}\big\{(\boldsymbol{X}\boldsymbol{\beta}_1 - \boldsymbol{X}\boldsymbol{\beta}_2)'(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})(\boldsymbol{X}\boldsymbol{\beta}_1 - \boldsymbol{X}\boldsymbol{\beta}_2) \\ & + 2\big[(\boldsymbol{X}\boldsymbol{\beta}_1 - \boldsymbol{X}\boldsymbol{\beta}_2)'(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})(\boldsymbol{X}\boldsymbol{\Gamma}_1 - \boldsymbol{X}\boldsymbol{\Gamma}_2)\boldsymbol{\mu}_x'\boldsymbol{\alpha} \\ & + \boldsymbol{\alpha}'(\boldsymbol{\mu_x\mu_x' + \Sigma_x})\boldsymbol{\alpha}\big((\boldsymbol{X}\boldsymbol{\Gamma}_1 - \boldsymbol{X}\boldsymbol{\Gamma}_2)\big)'(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})\big((\boldsymbol{X}\boldsymbol{\Gamma}_1 - \boldsymbol{X}\boldsymbol{\Gamma}_2)\big)\big\} \end{aligned} \tag{13}$$