Probability Model For A Multi-Predictor Linear Regression Model

Nan Wu
nanw@udel.edu

1 Linear Regression Model

The linear regression model is given as

$$y = \alpha + X\beta + \epsilon \tag{1}$$

where $y \in \mathbb{R}^N$, $X \in \mathbb{R}^{N \times P}$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}^P$, $\epsilon \in \mathbb{R}^N$, and $\epsilon \sim \text{Normal}\left(0, \sigma\right)$.
In this model, X is the predictor matrix, y is the outcome, α is the intercept, β is the coefficient vector associated with predictors, ϵ is the residual.

2 Probability Model

Stan allows us to use improper priors if we don't have any prior knowledge about the parameters. We can therefore start with a simple model by assuming improper priors for α, β and σ:

$$\begin{aligned}
\alpha &\sim \text{Uniform}\left(-\infty, \infty\right) \\
\beta_p &\sim \text{Uniform}\left(-\infty, \infty\right), p = 1, \cdots, P \\
\sigma &\sim \text{Uniform}\left(0, \infty\right)
\end{aligned}$$

Putting it all together, the probability model for the multi-predictor linear regression model is:

$$\begin{aligned}
y_n &\sim \text{Normal}\left(\mu_n, \sigma\right) \\
\mu_n &= \alpha + \sum_{p=1}^{P} X_{n,p}\beta_p \\
\alpha &\sim \text{Uniform}\left(-\infty, \infty\right) \\
\beta_p &\sim \text{Uniform}\left(-\infty, \infty\right), p = 1, \cdots, P \\
\sigma &\sim \text{Uniform}\left(0, \infty\right)
\end{aligned}$$