Probability Model For A Hierarchical Logistic Regression Model

Nan Wu
nanw@udel.edu

We consider a two level hierarchical logistic regression model in which the data are grouped into L distinct categories. Each category l gets its own coefficient vector β_l, and these coefficients are generated from the single set of hyperparameters.

1 Hierarchical Logistic Regression Model

The two level hierarchical logistic regression model is described as below:

$$\begin{aligned} y_n &\sim \text{Bernoulli}\left(p_n\right) \\ p_n &= \text{logit}^{-1}\left(\sum_{d=1}^{D} X_{n,d}\beta_{l_n,d}\right) \end{aligned}$$

where $y_n \in \{0,1\}, n = 1,\ldots,N$; $X \in \mathbb{R}^{N\times D}$ is the predictor matrix, where N is the number of observations, and D is the number of predictors; $l_n \in \{1,\ldots,L\}$ is the category associated with each binary outcome y_n; $\beta_l \in \mathbb{R}^D$ is the coefficient vector for category l.

2 Probability Model

$\beta_{l,d}$ are the parameters for this hierarchical logistic regression model, where $l \in \{1,\ldots,L\}, d = 1,\ldots,D$. Assume $\beta_{l,d}$ are drawn from a single set of hyperparameters for all categories l as below:

$$\begin{aligned} \beta_{l,d} &\sim \text{Normal}\left(\mu_{\text{d}}, \sigma_{\text{d}}\right) \\ d &= 1,\ldots,D \end{aligned}$$

Then we need to further specify the priors for the hyperparameters μ_d and σ_d. For simulation purpose, we use the following reasonable weakly informative priors for the location and scale parameters.

$$
\begin{aligned}
\mu &\sim \text{Normal}\,(0, 100) \\
\sigma &\sim \text{Gamma}\,(2, 0.1)
\end{aligned}
$$

where 0.1 is the rate parameter for gamma distribution.

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

$$
\begin{aligned}
y_n &\sim \text{Bernoulli}\,(p_n) \\
p_n &= \text{logit}^{-1}\left(\sum_{d=1}^{D} X_{n,d}\beta_{l_n,d}\right) \\
\beta_{l,d} &\sim \text{Normal}\,(\mu_{\text{d}}, \sigma_{\text{d}}) \\
\mu_d &\sim \text{Normal}\,(0, 100) \\
\sigma_d &\sim \text{Gamma}\,(2, 0.1)
\end{aligned}
$$

where $l_n \in (1, \ldots, L)$, and $d = 1, \ldots, D$.