STA 601 - Homework 15

Kedar Prabhudesai

November 13, 2013

Ordered Probit Model:

Model Specification:

X is $(p \times n)$ matrix of predictor variables and β is a $(p \times 1)$ vector of regression co-efficients. Y is $(n \times 1)$ vector of binary responses. We use an additional latent variable Z to define our ordered probit model.

$$\begin{aligned} y_i &\in \{1, 2, \ldots, d\}, i = 1, 2, \ldots, n \\ x_i &= (X_{i,1}, X_{i,2}, \ldots, X_{i,p}) \\ \beta &= (\beta_1, \beta_2, \ldots, \beta_p) \\ \pi_j(x_i) &= P(y_i = j \mid x_i, \beta) \end{aligned}$$

Adding latent variable,

$$\begin{aligned} y_i &= h(z_i, \tau) \\ z_i &\sim \mathcal{N}(x_i'\beta, 1) \end{aligned}$$

$h(z_i, \tau)$ is categorization function that maps z_i from real line to ordered values, $\{1, 2, \ldots, d\}$. Hence, $0 < \tau_2 < \tau_3 \ldots < \tau_{d-1}$

$$\begin{aligned} y_i = h(z_i, \tau) &= \sum_{j=3}^{d} j \times \mathbb{1}(z_i \in \{\tau_{j-1}, \tau_j\}) \\ \therefore \pi_j(x_i) &= P(z_i \in \{\tau_{j-1}, \tau_j\} \mid x_i, \beta) \\ &= P(z_i \leq \tau_j \mid x_i, \beta) - P(z_i \leq \tau_{j-1} \mid x_i, \beta) \\ &= \Phi(\tau_j - x_i'\beta) - \Phi(\tau_{j-1} - x_i'\beta) \end{aligned}$$

Where, Φ is cdf of standard normal distribution.

Likelihood:

$$L(x^n, y^n \mid \tau, \beta) = \prod_{j=1}^{d} \prod_{i=1}^{n} \pi_j(x_i)^{\mathbb{1}(y_i = j)}$$

Priors:

$$\begin{aligned}\beta &\sim \mathcal{N}_p(\beta_0, \Sigma_\beta)\\ \tau &\propto \mathbb{1}(0 < \tau_2 < \tau_3 \ldots < \tau_{d-1})\\ z_i &\sim \mathcal{N}(x_i'\beta, 1)\end{aligned}$$

Full Posterior:

$$p(\tau, \beta, z^n \mid x^n, y^n) \propto \mathcal{N}_p(\beta; \beta_0, \Sigma_\beta) \times \mathbb{1}(0 < \tau_2 < \tau_3 \ldots < \tau_{d-1}) \times \left[\prod_{i=1}^{n}\sum_{j=3}^{d} j \times \mathbb{1}(z_i \in \{\tau_{j-1}, \tau_j\})\right] \times \left[\prod_{i=1}^{n}\mathcal{N}(z_i; x_i'\beta, 1)\right]$$

Full Conditionals:

To compute this posterior we can use Gibbs Sampling, for which we need to compute full conditionals. In class we proved these,

$$p(\beta \mid x^n, y^n, z^n, \tau) \quad \propto \quad \mathcal{N}_p(\beta^*, \Sigma_\beta^*)$$

Where, (Referring to class notes,)

$$\Sigma_\beta^* \quad = \quad (\Sigma_\beta^{-1} + X'X)^{-1}$$

Let $\Sigma_\beta^{-1} = 0$. (Improper Prior)

$$\begin{aligned}\beta^* &= (X'X)^{-1}X'z\\ \Sigma_\beta^* &= (X'X)^{-1}\end{aligned}$$

$$p(z_i \mid z_{\sim i}, \beta, x^n, y^n, \tau) \quad \propto \quad \mathcal{N}(z_i; x_i'\beta, 1) \times \mathbb{1}(z_i \in \{\tau_{j-1}, \tau_j\})$$

Which is a normal distribution truncated between $\{\tau_{j-1}, \tau_j\}$.

$$p(\tau_j \mid \tau_{\sim j}, x^n, y^n, z^n, \beta) \quad \propto \quad \mathbb{1}(0 < \tau_2 < \tau_3 \ldots < \tau_{d-1}) \times \left[\prod_{i=1}^{n} j \times \mathbb{1}(z_i \in \{\tau_{j-1}, \tau_j\})\right]$$

Let us assume, that $\tau_j \propto \mathcal{N}(\mu_j, \sigma_j^2)$ constrained such that, $0 < \tau_2 < \tau_3 \ldots < \tau_{d-1}$. Further, let $a_j = max\{z_i : y_i = j-1\}$ and $b_j = min\{z_i : y_i = j\}$. The full conditional for τ_j will now be $\mathcal{N}(\mu_j, \sigma_j^2)$, truncated in the interval (a_j, b_j).

Specifying a prior on τ with given constraint can be a difficult task. As the number of categories b d goes up, prior with given constraints is even difficult. Further, the full conditional for τ_j becomes more truncated because the interval (a_j, b_j) becomes smaller. This will lead to a very sticky chain in the Gibbs Sampler. Hence, as sample size increases and number of categories increases leads to a sticky Gibbs Sampler.