

Summary of Short-term Research Objectives

Detian Deng

August 4, 2015

1 Model Specification

Let L be a K -dimensional Bernoulli random variable denoting the true state. Consider the general log linear model:

$$f(l; \Theta) = \exp\{\Theta_1^T l + \Theta_2^T u_2 + \dots + \Theta_K^T u_K - A^*(\Theta)\}$$

where U_k is a $\binom{K}{k} \times 1$ vector of k -way cross-products, $k = 1, \dots, K$, and $\Theta = (\Theta_1, \dots, \Theta_K)$ contains the natural parameters, which is a $(2^K - 1) \times 1$ vector.

Model restrictions, let $\tilde{l} = (l, u_2, \dots, u_K)^T$, and $S = \sum_{j=1}^K L_j = s$ has some fixed pmf

$$\begin{aligned} \pi(s) &:= P(S = s) \\ &= \frac{1}{A(\Theta)} \sum_{\tilde{l}: S=s} \exp\{\Theta^T \tilde{l}\}, \quad s = 0, 1, \dots, K \end{aligned} \tag{1}$$

$$A(\Theta) = \sum_{\tilde{l}: l \in \{0,1\}^K} \exp\{\Theta^T \tilde{l}\} \tag{2}$$

2 Posterior Distribution

2.1 General Model

For General model, \tilde{L} is a square matrix with dimension $J_1 = 2^K - 1$. Recall that

$$A(\Theta) = \frac{1}{\pi(0)}$$

$$\pi(s) = \frac{1}{A(\Theta)} \sum_{\tilde{l}: S=s} \exp\{\Theta^T \tilde{l}\}, s = 1, \dots, K \quad (3)$$

$$\mu_k = \frac{1}{A(\Theta)} \sum_{\tilde{l}: l_k=1} \exp\{\Theta^T \tilde{l}\}, k = 1, \dots, K \quad (4)$$

Define intermediate parameter $\phi_j = \exp(\theta^T \tilde{l}_j) > 0$, $\Theta = (\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(K)})$, $j = 1, \dots, J_1$ and two $K \times J_1$ sub-design matrices B, C , where $B[k, j] = 1(\sum_{s=1}^K \tilde{L}[j, s] = k)$, $C[k, j] = \tilde{L}[j, k]$. Thus (3) and (4) become

$$\vec{\phi} > 0 \quad (5)$$

$$B\vec{\phi} = \vec{\pi}/\pi(0) \quad (6)$$

$$C\vec{\phi} = \vec{\mu}/\pi(0) \quad (7)$$

Note that B and C are not independent constraints and should be compatible so that $\begin{pmatrix} B \\ C \end{pmatrix}$ has rank $2K - 1$.

Based on [1][2], we can sample $\vec{\phi}$ from Uniform distribution subject to the above linear constraints efficiently and robustly. Then Θ are the solutions to the linear system (J_1 equations with J_1 unknowns):

$$\tilde{L}\Theta = \log \vec{\phi}$$

Then the posterior distribution of (μ, π) given data L becomes:

$$\begin{aligned} P(\mu, \pi | L) &\propto P(L, \mu, \pi) \\ &= P(L | \mu, \pi) P(\mu, \pi) \\ &= \int P(L, \phi | \mu, \pi) d\phi P(\mu, \pi) \\ &\propto \int P(L | \phi) P(\phi | \mu, \pi) d\phi P(\mu, \pi) \\ &\approx \sum P(L | \phi_h) P(\pi | \mu) P(\mu) \end{aligned}$$

where $P(L | \phi_h)$ is the log linear model density evaluated at ϕ_h , which is sampled uniformly from the feasible region subject to constraints (5-7).

$P(\mu)$ is a multivariate logit Normal density with mean 0 and standard deviation Σ .

$P(\pi | \mu) = P(\pi_0 | \mu) P(\pi_1, \dots, \pi_K | \pi_0, \alpha)$, where $P(\pi_0 | \mu)$ is a uniform density on $(0, 1 - \max(\mu))$ and $P(\pi_1, \dots, \pi_K | \pi_0, \alpha)$ is a Stick Breaking prior with shape parameter α on $(0, 1 - \pi_0)$.

2.2 Sampling Algorithm

Algorithm 1 Usample(n, μ, π_0, π_{0-})

- 1: **if** neither μ, π_{0-} is NULL **then**
 - 2: sample ϕ uniformly from feasible region defined by

$$\begin{aligned} \phi &> 0 \\ B\phi &= \pi_{0-} / \pi_0 \\ C\phi &= \mu / \pi_0 \end{aligned}$$
 - 3: **else if** only μ is NULL **then**
 - 4: sample ϕ uniformly from feasible region defined by

$$\begin{aligned} \phi &> 0 \\ B\phi &= \pi_{0-} / \pi_0 \end{aligned}$$
 - 5: **else**
 - 6: sample ϕ uniformly from feasible region defined by

$$\begin{aligned} \phi &> 0 \\ C\phi &= \mu / \pi_0 \end{aligned}$$
 - 7: **return** n samples of ϕ
-

Algorithm 2 SampleByBlock($n_{Iter}, n_{Burn}, L, \mu_{init}, \pi_{init}, \sigma$)

- 1: Initialize $\mu^{(0)} = \mu_{init}, \pi^{(0)} = \pi_{init}$
 - 2: **for** $t \in 1$ to $(n_{Iter} + n_{Burn})$ **do**
 - 3: $\phi = \text{Usample}(1, \mu = \text{NULL}, (\pi_0, \pi_{0-}) = \pi^{(t-1)})$
 - 4: $\mu^* = \pi_0^{(t-1)} C\phi$
 - 5: $\alpha = \frac{P(L, \mu^*, \pi^{(t-1)})}{P(L, \mu^{(t-1)}, \pi^{(t-1)})}$
 - 6: $\mu^{(t)} = \mu^*$ with probability $\max(1, \alpha)$; other wise, $\mu^{(t)} = \mu^{(t-1)}$
 - 7: $\pi_0^* = \text{logit}^{-1}[\text{logit}(\pi_0^{(t-1)}) + \epsilon]$, where $\epsilon \sim N(0, \sigma)$.
 - 8: $\phi = \text{Usample}(1, \mu = \mu^{(t)}, \pi_0 = \pi_0^*, \pi_{0-} = \text{NULL})$
 - 9: $\pi_{0-}^* = \pi_0^* B\phi$
 - 10: $\pi^* = (\pi_0^*, \pi_{0-}^*)$
 - 11: $\alpha = \frac{P(L, \mu^{(t)}, \pi^*)}{P(L, \mu^{(t)}, \pi^{(t-1)})}$
 - 12: $\pi^{(t)} = \pi^*$ with probability $\max(1, \alpha)$; other wise, $\pi^{(t)} = \pi^{(t-1)}$
 - 13: **return** $\mu^{(n_{Burn}+1)}$ to $\mu^{(n_{Burn}+n_{Iter})}$ and $\pi^{(n_{Burn}+1)}$ to $\pi^{(n_{Burn}+n_{Iter})}$
-

2.3 Measurement Error Model

Reference

1. Smith RL . Efficient Monte-Carlo Procedures for Generating Points Uniformly Dis- tributed over Bounded Regions. Operations Research, 32(6), 12961308 (1984).
2. Van den Meersche, Karel, K. E. R. Soetaert, and D. J. Van Oevelen. "xsample (): an R function for sampling linear inverse problems." Journal of Statistical Software 30 (2009).