Summary of Short-term Research Objectives

Detian Deng

September 3, 2015

1 Log Linear Model Representation

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

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

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

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

$$\begin{aligned}\pi(s) :=& P(S = s) \\ =& \frac{1}{A(\Theta)} \sum_{\tilde{l}:S=s} \exp\{\Theta^T \tilde{l}\}\ , s = 0, 1, \ldots, K \qquad (1)\\ A(\Theta) =& \sum_{\tilde{l}:l\in\{0,1\}^K} \exp\{\Theta^T \tilde{l}\} \qquad (2)\end{aligned}$$

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

$$\begin{aligned}A(\Theta) =& \frac{1}{\pi(0)} \\ \pi(s) =& \frac{1}{A(\Theta)} \sum_{\tilde{l}:S=s} \exp\{\Theta^T \tilde{l}\}\ , s = 1, \ldots, K \qquad (3)\\ \mu_k =& \frac{1}{A(\Theta)} \sum_{\tilde{l}:l_k=1} \exp\{\Theta^T \tilde{l}\}\ , k = 1, \ldots, K \qquad (4)\end{aligned}$$

Define intermediate parameter $\phi_j = \exp(\theta^T \tilde{l}_j) > 0$, $\Theta = (\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(K)}), j = 1, \ldots, J_1$ 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

$$\begin{aligned}\vec{\phi} >& 0 \qquad (5)\\ B\vec{\phi} =& \vec{\pi}/\pi(0) \qquad (6)\\ C\vec{\phi} =& \vec{\mu}/\pi(0) \qquad (7)\end{aligned}$$

Note that B and C are not independent constraints and should be compatible so that $\binom{B}{C}$ has rank $2K - 1$. Explicitly, μ and π must satisfy

$$\sum_{k=1}^K \mu_k = \sum_{k=1}^K k\pi_k$$

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):

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

2 Multinomial Representation

Considering all possible outcomes of the K-dimensional Bernoulli random vector as a 2^K-dimensional Multinomial random vector, if all the Multinomial cell probabilities are strictly positive, then there is a a one-to-one mapping between the cell probabilities p and the canonical parameters Θ in the saturated log linear model through the intermediate parameters ϕ defined above.

$$\begin{aligned} p_0 =& \pi_0 = \frac{1}{1+\sum \phi_j} \\ p_j =& \pi_0 \phi_j \end{aligned}$$

2.1 Sparsity of Multinomial Model

If we allow some of the cell probabilities to be exactly 0, then the model violates the log linear model framework, but gains useful model sparsity to our interests and the above relations among μ, π, ϕ and p still hold.

In the etiology study, scientists have strong beliefs that there cannot be too many (like more than 3) pathogens that jointly cause the disease. In other words, for some fixed integer $1 \leq S_{max} \leq K$, P($S > S_{max}$) = 0

3 Joint Density

3.1 Prior on (μ, π)

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 constrains (5-7).
$P(\mu)$ is a multivariate logit Normal density with mean 0 and standard deviation Σ.
$P(\pi|\mu) = P(\pi_0|\mu) P(\pi_1, .., \pi_K | \pi_0, \alpha)$, where $P(\pi_0|\mu)$ is a uniform density on $(0, 1 - \max(\mu))$ and $P(\pi_1, ..\pi_K | \pi_0, \alpha)$ is a Stick Breaking prior with shape parameter α on $(0, 1 - \pi_0)$.

3.1.1 Sampling Algorithm

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

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

$$\begin{aligned} \phi >& 0 \\ B\phi =& \pi_{0-}/\pi_0 \\ C\phi =& \mu/\pi_0 \end{aligned}$$

else if only μ is NULL **then**
    sample ϕ uniformly from feasible region defined by

$$\begin{aligned} \phi >& 0 \\ B\phi =& \pi_{0-}/\pi_0 \end{aligned}$$

else
    sample ϕ uniformly from feasible region defined by

$$\begin{aligned} \phi >& 0 \\ C\phi =& \mu/\pi_0 \end{aligned}$$

return n samples of ϕ

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

Initialize $\mu^{(0)} = \mu_{init}$, $\pi^{(0)} = \pi_{init}$
for $t \in 1$ to $(n_{Iter} + n_{Burn})$ **do**
    $\phi =$ Usample$(1, \mu =$ NULL$, (\pi_0, \pi_{0-}) = \pi^{(t-1)})$
    $\mu^* = \pi_0^{(t-1)} C\phi$
    $\alpha = \frac{P(L,\mu^*,\pi^{(t-1)})}{P(L,\mu^{(t-1)},\pi^{(t-1)})}$
    $\mu^{(t)} = \mu^*$ with probability $\max(1, \alpha)$; other wise, $\mu^{(t)} = \mu^{(t-1)}$
    $\pi_0^* =$ logit$^{-1}[logit(pi_0^{(t-1)}) + \epsilon]$, where $\epsilon \sim$ N$(0, \sigma)$.
    $\phi =$ Usample$(1, \mu = \mu^{(t)}, \pi_0 = \pi_0^*, \pi_{0-} =$ NULL$)$
    $\pi_{0-}^* = \pi_0^* B\phi$
    $\pi^* = (\pi_0^*, \pi_{0-}^*)$
    $\alpha = \frac{P(L,\mu^{(t)},\pi^*)}{P(L,\mu^{(t)},\pi^{(t-1)})}$
    $\pi^{(t)} = \pi^*$ with probability $\max(1, \alpha)$; other wise, $\pi^{(t)} = \pi^{(t-1)}$
return $\mu^{(n_{Burn}+1)}$ to $\mu^{(n_{Burn}+n_{Iter})}$ and $\pi^{(n_{Burn}+1)}$ to $\pi^{(n_{Burn}+n_{Iter})}$

3.2 Measurement Error Model

Let M be the observed measurement, and L be the latent true state, with γ and δ representing the TPR and FPR.

$$\begin{aligned} P(\mu, \pi, \gamma, \delta | M) \propto & P(M, \mu, \pi, \gamma, \delta) \\ = & P(M | \mu, \pi, \gamma, \delta) P(\mu, \pi, \gamma, \delta) \\ = & \sum_{L \in \mathbb{L}} P(M, L | \mu, \pi, \gamma, \delta) P(\mu, \pi, \gamma, \delta) \\ = & \sum_{L \in \mathbb{L}} \big[P(M | L, \gamma, \delta) P(L | \mu, \pi) \big] P(\mu, \pi) P(\gamma) P(\delta) \end{aligned}$$

where

$$\begin{aligned} P(M | L, \gamma, \delta) &= \prod_{k=1}^{K} P(M_k | L_k, \gamma, \delta) \text{ (conditional independence assumption)} \\ &= \prod_{k=1}^{K} (\gamma^{L_k} \delta^{1-L_k})^{M_k} [(1-\gamma)^{L_k} (1-\delta)^{1-L_k}]^{1-M_k} \end{aligned}$$

$P(\gamma)$ and $P(\delta)$ are beta priors, and $P(L|\mu, \pi)P(\mu, \pi)$ is defined in the same way as in previous section.

3.2.1 Sampling Algorithm

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).