Summary of Short-term Research Objectives

Detian Deng

June 1, 2015

1 Model Specification

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

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

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

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

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

2 Posterior Distribution

$$\begin{aligned}
P(\mu,\theta^{(2)}|L) \propto & P(L,\mu,\theta^{(2)}) \\
\propto & P(L,\mu,\theta^{(1)},\theta^{(2)},\pi) \\
\propto & P(L|\mu,\theta^{(1)},\theta^{(2)},\pi)P(\mu,\theta^{(1)},\theta^{(2)},\pi) \\
\propto & P(L|\theta^{(1)},\theta^{(2)})P(\theta^{(1)},\theta^{(2)}|\mu,\pi)P(\mu)P(\pi) \\
\propto & \mathrm{QE}(L;\theta^{(1)},\theta^{(2)})\mathrm{UFR}(\theta^{(1)},\theta^{(2)}|\mu,\pi)\mathrm{N}(\mathrm{logit}(\mu),\Sigma)\mathrm{tPois}(\pi)
\end{aligned}$$

where QE is the second-order log linear model.
UFR is a Multivariate distribution of $[\theta^{(1)},\theta^{(2)}|\mu,\pi]$ subject to non-linear constrains: $M(\theta^{(1)},\theta^{(2)})=\mu$ and $\Pi(\theta^{(1)},\theta^{(2)})=\pi$, which can be sampled by a two-step procedure.
tPois is a truncated conjugate Poisson distribution defined as:

$$\begin{aligned}
\pi \sim & \mathrm{Dirichilet}(\mathrm{hist}(\vec{s})) \\
s \sim & \frac{\lambda^s}{s!}e^{-\lambda}/[1-\sum_{s>K}\frac{\lambda^s}{s!}e^{-\lambda}]
\end{aligned}$$

2.1 On sampling $[\Theta|\mu,\pi]$

2.1.1 General model

For General 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 $B[k,j]=1(\tilde{l}_j^T1=k)$, $C[k,j]=1(\tilde{L}[j,k]=1)$. Thus (6) and (7) become

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

Note that B and C are not independent constraints and should be compatible so that $\binom{B}{C}$ 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):

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

Since $\vec{\phi}$ fully specifies all cell probabilities, the posterior distribution becomes

$$P(\mu,\vec{\phi}|L)\propto \mathrm{LL}(L;\vec{\phi})\mathrm{UFR}(\vec{\phi}|\mu,\pi)\mathrm{N}(\mathrm{logit}(\mu),\Sigma)\mathrm{tPois}(\pi)$$

2.1.2 QE model

For QE model, $\tilde{L}$ is a $J_1 \times J_2$, where $J_1 = 2^K - 1, J_2 = \frac{1}{2}K(K+1)$. Let $\theta = (\theta^{(1)}, \theta^{(2)})$, we have a over-determined linear system after sampling $\vec{\phi}$: (J_1 equations with J_2 unknowns)

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

In this case, we can first find Θ as in last section and then solve the constrained Least Square problem to get the QE parameter estimates θ:

$$\begin{aligned} \text{minimize } & ||\theta - \Theta||_2^2 \\ \text{subject to } & Be^{\tilde{L}\theta} = \vec{\pi}/\pi(0) \\ \text{and } & Ce^{\tilde{L}\theta} = \vec{\mu}/\pi(0) \end{aligned}$$

2.2 Example for $K = 3$

For $K = 3$ with the General Log-linear model, $\Theta = (\theta_1, \theta_2, \theta_3, \theta_{12}, \theta_{13}, \theta_{23}, \theta_{123})^T$
$\Phi = (\phi_1 = \exp(\theta_1), \ldots, \phi_4 = \exp(\theta_1 + \theta_2 + \theta_{12}), \ldots, \phi_7 = \exp(\theta_1 + \theta_2 + \theta_3 + \theta_{12} + \theta_{13} + \theta_{23} + \theta_{123}))^T$
When we made prior assumption on $\vec{\mu} = (\mu_1, \mu_2, \mu_3)^T$ and π_0, the sampling constraints are:

$$\begin{aligned}
\sum_{i=1}^{7} \phi_i =& \frac{1-\pi_0}{\pi_0} \\
\phi_1 + \phi_4 + \phi_5 + \phi_7 =& \frac{\mu_1}{\pi_0} \\
\phi_2 + \phi_4 + \phi_6 + \phi_7 =& \frac{\mu_2}{\pi_0} \\
\phi_3 + \phi_5 + \phi_6 + \phi_7 =& \frac{\mu_3}{\pi_0} \\
\phi_i \in& (0, \frac{1}{\pi_0} - 1)
\end{aligned}$$

And an implicit assumption is $\mu_k < 1 - \pi_0 \leq \sum_{k=1}^{3} \mu_k$.

By applying row reduction to the augmented matrix of the above 4 equations, we have

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & -1 & -1 & \frac{\mu_2+\mu_3+\pi_0-1}{\pi_0} \\
0 & 1 & 0 & 0 & -1 & 0 & -1 & \frac{\mu_1+\mu_3+\pi_0-1}{\pi_0} \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & -\frac{\mu_3}{\pi_0} \\
0 & 0 & 0 & 1 & 1 & 1 & 2 & \frac{1-\pi_0-\sum_{k=1}^{3}\mu_k}{\pi_0}
\end{pmatrix}$$

Thus let $\phi_5 = u, \phi_6 = v, \phi_7 = w$, the feasible region of Φ is defined by:

$$\begin{pmatrix}
v + w - \frac{\mu_2+\mu_3+\pi_0-1}{\pi_0} \\
u + w - \frac{\mu_1+\mu_3+\pi_0-1}{\pi_0} \\
-u - v - w + \frac{\mu_3}{\pi_0} \\
-u - v - 2w - \frac{1-\pi_0-\sum_{k=1}^{3}\mu_k}{\pi_0} \\
u \\
v \\
w
\end{pmatrix}
=
\begin{pmatrix}
-\frac{\mu_2+\mu_3+\pi_0-1}{\pi_0} \\
-\frac{\mu_1+\mu_3+\pi_0-1}{\pi_0} \\
\frac{\mu_3}{\pi_0} \\
-\frac{1-\pi_0-\sum_{k=1}^{3}\mu_k}{\pi_0} \\
0 \\
0 \\
0
\end{pmatrix}
+ u \begin{pmatrix} 0 \\ 1 \\ -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}
+ v \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}
+ w \begin{pmatrix} 1 \\ 1 \\ -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

and $\phi_i \in (0, \frac{1}{\pi_0} - 1)$ for $i = 1, \ldots, 7$

To sample (u, v, w) from its feasible space, we can do as follow:

1. Sample v from Unif$(0, \frac{1-\mu_1-\pi_0}{\pi_0})$.
2. Sample u from Unif$(0, \frac{\mu_3}{\pi_0} - v)$.
3. Sample w from Unif$(\max(\frac{\mu_2+\mu_3+\pi_0-1}{\pi_0} - v, \frac{\mu_1+\mu_3+\pi_0-1}{\pi_0} - u), \frac{\mu_3}{\pi_0} - v - u)$.

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).