Supplemental material to "Dirichlet Process Mixture Models using Matrix-Generalized Half-t Distribution"

January 28, 2023

Sanghyun Lee[1], Chanmin Kim[1,2*]
[1]Department of Statistics, Sungkyunkwan University, Seoul, Korea
[2]Department of Biostatistics, Boston University, Boston, MA
**email*: chanmin.kim@skku.edu

1 CDP

- Posteriors in each Gibbs Sampling Step

1. Update $\boldsymbol{\mu}_j$.

$$p\left(\boldsymbol{\mu}_j \mid -\right) \sim \mathcal{N}_p\left((n_j+\rho)^{-1}\left(\sum_{i:c_i=j} \boldsymbol{x}_i + \rho\boldsymbol{\xi}\right),\ \left((n_j+\rho)\, S_j\right)^{-1}\right)$$

2. Update S_j.

$$p\left(S_j \mid -\right) \sim \mathcal{W}_p\left(n_j+\beta+1,\ \left(\sum_{i:c_i=j}\left(\boldsymbol{x}_i-\boldsymbol{\mu}_j\right)\left(\boldsymbol{x}_i-\boldsymbol{\mu}_j\right)' + \rho\left(\boldsymbol{\mu}_j-\boldsymbol{\xi}\right)\left(\boldsymbol{\mu}_j-\boldsymbol{\xi}\right)' + \beta W\right)^{-1}\right)$$

3. Update $\boldsymbol{\xi}$.

$$p\left(\boldsymbol{\xi} \mid -\right) \sim \mathcal{N}_p\left(\left(\rho\sum_{j=1}^{k} S_j + \Sigma_y^{-1}\right)^{-1}\Sigma_y^{-1}\boldsymbol{\mu}_y,\ \left(\rho\sum_{j=1}^{k} S_j + \Sigma_y^{-1}\right)^{-1}\right)$$

4. Update ρ.

$$p\left(\rho \mid -\right) \sim \mathcal{G}\left(\frac{pk}{2}+\frac{1}{4},\ \left(\frac{1}{2}+\frac{1}{2}\sum_{j=1}^{k}\left(\boldsymbol{\mu}_j-\boldsymbol{\xi}\right)' S_j\left(\boldsymbol{\mu}_j-\boldsymbol{\xi}\right)\right)^{-1}\right)$$

5. Update W.

$$p\left(W \mid -\right) \sim \mathcal{W}_p\left(k\beta+p,\ \left(\beta\sum_{j=1}^{k} S_j + p\Sigma_y^{-1}\right)^{-1}\right)$$

6. Update β using ARMS.

$$\log p\left(\beta^* \mid -\right) \;\propto\; \frac{pk\left(1/\beta^*+p-1\right)}{2}\log\left(\frac{\left(1/\beta^*+p-1\right)}{2}\right)+\frac{k\left(1/\beta^*+p-1\right)}{2}\log\left(|W|\right)- k\log\Gamma_p\left(\frac{\left(1/\beta^*+p-1\right)}{2}\right)+\frac{1/\beta^*-2}{2}\sum_{j=1}^{k}\log|S_j|-\frac{\left(1/\beta^*+p-1\right)}{2}\operatorname{tr}\left(W\sum_{i=1}^{k}S_j\right)- \frac{1}{2}\log\beta^*-\frac{p}{2}\beta^* \text{ where } \beta^*=\frac{1}{\beta-p+1}$$

7. Update component membership indicator c_i.

$$\log p\left(\boldsymbol{x}_i \mid c_{-i},\ \boldsymbol{\xi},\ \rho,\ W,\ \beta\right) \;=\; \frac{p}{2}\log\frac{n_{-i,j}+\rho}{n_{-i,j}+\rho+1}-\frac{p}{2}\log\pi+\log\Gamma\left(\frac{n_{-i,j}+\beta+1}{2}\right)- \log\Gamma\left(\frac{n_{-i,j}+\beta+1-p}{2}\right)+\frac{n_{-i,j+\beta}}{2}\log|W^*|- \frac{n_{-i,j}+\beta+1}{2}\log\left|W^*+\frac{n_{-i,j}+\rho}{n_{-i,j}+\rho+1}\left(\boldsymbol{x}_i-\boldsymbol{\xi}^*\right)\left(\boldsymbol{x}_i-\boldsymbol{\xi}^*\right)'\right| \text{ where}$$

$$W^*=\beta W+\rho\boldsymbol{\xi}\boldsymbol{\xi}'+\sum_{l:c_l=j}\boldsymbol{x}_l\boldsymbol{x}_l'-(\rho+n_{-i,j})\boldsymbol{\xi}^*\boldsymbol{\xi}^{*'} \text{ and } \boldsymbol{\xi}^*=\left(\rho\boldsymbol{\xi}+\sum_{l:c_l=j}\boldsymbol{x}_l\right)/(\rho+n_{-i,j})$$

8. Update α using ARMS.

$$\log p\left(\alpha^* \mid -\right) \;\propto\; \left(-k-\frac{1}{2}\right)\log\alpha^*+\log\Gamma\left((\alpha^*)^{-1}\right)-\log\Gamma\left(n+(\alpha^*)^{-1}\right)-\frac{\alpha^*}{2}$$

where $\alpha^*=1/\alpha$

2 CCDP

- Priors

$$
\begin{aligned}
f\left(\boldsymbol{x}_i \mid \boldsymbol{\mu}_i,\ S_i\right) &\sim \mathcal{N}_p\left(\boldsymbol{\mu}_i,\ S_i^{-1}\right) \text{ where } S_i \text{ is precision.}\\[1em]
\left(\boldsymbol{\mu}_j \mid \boldsymbol{\xi},\ R\right) &\sim \mathcal{N}_p\left(\boldsymbol{\xi},\ R^{-1}\right) \text{ where } R \text{ is precision.}\\
\boldsymbol{\xi} &\sim \mathcal{N}_p\left(\boldsymbol{\mu}_y,\ \Sigma_y\right)\\
R &\sim \mathcal{W}_p\left(p,\ (p\Sigma_y)^{-1}\right)\\[2em]
\left(S_j \mid \beta,\ W\right) &\sim \mathcal{W}_p\left(\beta,\ (\beta W)^{-1}\right)\\
W &\sim \mathcal{W}_p\left(p,\ \Sigma_y/p\right)\\
\frac{1}{\beta-p+1} &\sim \mathcal{G}\left(\frac{1}{2},\ \frac{2}{p}\right)\\[2em]
\alpha^{-1} &\sim \mathcal{G}\left(\frac{1}{2},\ 2\right)
\end{aligned}
$$

- Posteriors in each Gibbs Sampling Step

1. Update $\boldsymbol{\mu}_j$.

$$p\left(\boldsymbol{\mu}_j \mid -\right) \ \sim \ \mathcal{N}_p\left(\left(n_j S_j + R\right)^{-1}\left(S_j \sum_{i:c_i=j} \boldsymbol{x}_i + R\boldsymbol{\xi}\right),\ \left(n_j S_j + R\right)^{-1}\right)$$

2. Update S_j.

$$p\left(S_j \mid -\right) \ \sim \ \mathcal{W}_p\left(n_j + \beta,\ \left(\sum_{i:c_i=j} \left(\boldsymbol{x}_i - \boldsymbol{\mu}_j\right)\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right)' + \beta W\right)^{-1}\right)$$

3. Update $\boldsymbol{\xi}$.

$$p\left(\boldsymbol{\xi} \mid -\right) \ \sim \ \mathcal{N}_p\left(\left(kR + \Sigma_y^{-1}\right)^{-1}\left(\sum_{j=1}^{k} R\boldsymbol{\mu}_j + \Sigma_y^{-1}\boldsymbol{\mu}_y\right), \left(kR + \Sigma_y^{-1}\right)^{-1}\right)$$

4. Update R.

$$p\left(R \mid -\right) \ \sim \ \mathcal{W}_p\left(k + p,\ \left(\sum_{j=1}^{k} \left(\boldsymbol{\mu}_j - \boldsymbol{\xi}\right)\left(\boldsymbol{\mu}_j - \boldsymbol{\xi}\right)' + p\Sigma_y\right)^{-1}\right)$$

5. Update W.

$$p\left(W \mid -\right) \ \sim \ \mathcal{W}_p\left(k\beta + p,\ \left(\beta \sum_{j=1}^{k} S_j + p\Sigma_y^{-1}\right)^{-1}\right)$$

6. Update β using the same method mentioned above in CDP case.
7. Update component membership indicator c_i.
   - When $n_{-i,j} > 0,\ p\left(c_i = j \mid -\right) = \dfrac{n_{-i,j}}{n - 1 + \alpha}\mathcal{N}_p\left(\boldsymbol{x}_i \mid \boldsymbol{\mu}_j,\ S_j^{-1}\right),$
   - When $n_{-i,j} = 0,\ p\left(c_i \neq c_{i'} \text{ where } i \neq i' \mid -\right) = \dfrac{\alpha}{n - 1 + \alpha}\displaystyle\int p\left(\boldsymbol{x}_i \mid \boldsymbol{\mu},\ S\right) p(\boldsymbol{\mu},\ S \mid \boldsymbol{\xi},\ R,\ W)\, \mathrm{d}\boldsymbol{\mu}\, \mathrm{d}S$
8. Update α using the same method mentioned above in CDP case.