Double Collapsed Gibbs sampler

- Goal: Sample $[S \mid Y]$, $[Y_{n+1} \mid Y]$
- Idea: use gibbs sampler, conditionals
  $[s_i \mid s_{-i}, Y]$
- Derivation:
  Note: $[S, Y] = [S, Y_i \mid Y_{-i}][Y_i]$

  so $[s_i \mid s_{-i}, Y] \overset{s_i}{\propto} [S, Y]$
  $\overset{s_i}{\propto} [s_i, Y_i \mid S_{-i}, Y_{-i}]$
- Let $Y_{-i,j} = (Y_\ell : \ell \neq i, s_\ell = j)$
  so $[s_i = j \mid s_{-i}, Y] \overset{s_i}{\propto} [Y_i \mid s_i = j, s_{-i}, Y_{-i}][s_i = j \mid s_{-i}, Y_{-i}]$
  $= [Y_i \mid s_i = j, Y_{-i,j}][s_i = j \mid s_{-i}]$

Let $k = \#$ unique s_ℓ s.t. $\ell \in -i$
$n_{-i,j} = \# \ \ell \in -i$ s.t. $s_\ell = j$

Thm $P(s_i = j \mid s_{-i}) = \begin{cases} \dfrac{\alpha}{\alpha + n - 1} & j = k+1 \\ \dfrac{n_{-i,j}}{\alpha + n - 1} & j \in \{1, \dots, k\} \end{cases}$

let θ_j be param of j^{th} cluster

Then $[y_i \mid s_i = j, y_{-i,j}] = \int [y_i \mid s_i = j, y_{-i,j}, \theta_j][\theta_j \mid s_i = j, y_{-i,j}] d\theta_j$

$= \int N(y_i; \theta_j) [\theta_j \mid s_i = j, y_{-i,j}] d\theta_j$

& $[\theta_j \mid s_i = j, y_{-i,j}] \overset{\theta_j}{\propto} [y_{-i,j} \mid s_i = j, \theta_j][\theta_j \mid s_i = j]$

$= [y_{-i,j} \mid s_i = j, \theta_j][\theta_j]$

$= \left(\prod_{\substack{l : s_l = j \\ l \neq i}} N(y_l; \theta_j) \right) g_0(\theta_j)$

So $[y_i \mid s_i = j, y_{-i,j}] = \int N(y_i, \theta_j) \left(\prod_{\substack{l : s_l = j \\ l \neq i}} N(y_l; \theta_j) \right) g_0(\theta_j) d\theta_j$

$\theta_j = (\mu_j, \tau_j^{-1}) \sim$ normal gamma $(\mu_0, k, a_\tau, b_\tau)$

so $[y_i \mid s_i = j, y_{-i,j}] \propto$ posterior for normal likelihood normal-gamma prior

- $[Y_{n+1} \mid Y]$

Sample $S^{(t)} \sim [S \mid Y]$ $t=1,\dots,N$ using gibbs sampler.

Let $k^{(t)} = \#$ unique $S_\ell^{(t)}$ s.t. $\ell \in \{1,\dots,n\}$

$n_j^{(t)} = \#\ \ell \in \{1,\dots,n\}$ s.t. $S_\ell^{(t)} = j$ $(j \in \{1,\dots,k\})$

$Y_{(j)}^{(t)} = (Y_\ell : \ell \in \{1,\dots,n\},\ S_\ell^{(t)} = j)$ $(j \in \{1,\dots,k\})$

Then from gibbs sampler derivations,

$[Y_{n+1} \mid S_{n+1} = j, Y_{(j)}^{(t)}]$ = posterior rng$(Y_\ell : S_\ell^{(t)} = j$, $\ell \in \{1,\dots,n\})$

$$P(S_{n+1} = j \mid S^{(t)}) = \begin{cases} \dfrac{\alpha}{\alpha+n-1} & j = k+1 \\ \dfrac{n_j^{(t)}}{\alpha+n-1} & j \in \{1,\dots,k\} \end{cases}$$

$$\therefore [Y_{n+1} \mid Y] = \int\int [Y_{n+1} \mid S_{n+1}, S, Y][S_{n+1} \mid S, Y][S \mid Y]\, dS_{n+1}$$

$$\approx \frac{1}{N}\sum_{t=1}^{N} \int [Y_{n+1} \mid S_{n+1}, S^{(t)}, Y][S_{n+1} \mid S^{(t)}, Y]\, dS_{n+1}$$

$$= \frac{1}{N}\sum_{t=1}^{N}\sum_{j=1}^{k+1} [Y_{n+1} \mid S_{n+1} = j, Y_{(j)}^{(t)}][S_{n+1} = j \mid S^{(t)}]$$