Stat 289: Bayesian Methods and Computation
Lecture 8: Additional Notes

1. Gibbs Sampling for Mixture of Normals Model

The mixture-of-normals model is

$$\begin{aligned}
y_i|z_i = j, \boldsymbol{\mu}, \boldsymbol{\sigma}^2 &\sim \text{Normal}(\mu_j, \sigma_j^2), \quad (j = 0 \text{ or } 1) \\
z_i|\alpha &\sim \text{Bernoulli}(\alpha) \\
p(\mu_0, \sigma_0^2) &\propto (\sigma_0^2)^{-1} \\
p(\mu_1, \sigma_1^2) &\propto (\sigma_1^2)^{-1} \\
p(\alpha) &\propto 1
\end{aligned}$$

The joint posterior distribution for the parameters and indicator variables is

$$\begin{aligned}
p(\alpha, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{z}|\mathbf{y}) &\propto p(\alpha)\, p(\mu_0, \sigma_0^2)\, p(\mu_1, \sigma_1^2)\, p(\mathbf{z}|\alpha)\, p(\mathbf{y}|\mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) \\
&\propto (\sigma_0^2)^{-1}\, (\sigma_1^2)^{-1} \left[\prod_{i=1}^{n} p(z_i|\alpha)\right] \left[\prod_{i=1}^{n} p(y_i|z_i, \boldsymbol{\mu}, \boldsymbol{\sigma}^2)\right]
\end{aligned}$$

The Gibbs sampling algorithm draws from the following distributions:

1. Sample z_i from $p(z_i|\alpha, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{y})$ for $i = 1, \ldots, n$
2. Sample α from $p(\alpha|\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y})$
3. Sample μ_0 from $p(\mu_0|\alpha, \mu_1, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y})$
4. Sample μ_1 from $p(\mu_1|\alpha, \mu_0, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y})$
5. Sample σ_0^2 from $p(\sigma_0^2|\alpha, \boldsymbol{\mu}, \sigma_1^2, \mathbf{z}, \mathbf{y})$
6. Sample σ_1^2 from $p(\sigma_1^2|\alpha, \boldsymbol{\mu}, \sigma_0^2, \mathbf{z}, \mathbf{y})$

The full conditional distributions are derived below.

1. Sampling the indicators $z_i,\ i = 1, \ldots, n$.

$$\begin{aligned}
Pr(z_i = 1|\alpha, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{y}) &= \frac{Pr(z_i = 1|\alpha)p(y_i|z_i = 1, \boldsymbol{\mu}, \boldsymbol{\sigma}^2)}{Pr(z_i = 1|\alpha)p(y_i|z_i = 1, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) + Pr(z_i = 0|\alpha)p(y_i|z_i = 0, \boldsymbol{\mu}, \boldsymbol{\sigma}^2)} \\
&= \frac{\alpha N(y_i|\mu_1, \sigma_1^2)}{\alpha N(y_i|\mu_1, \sigma_1^2) + (1-\alpha)N(y_i|\mu_0, \sigma_0^2)}
\end{aligned}$$

For Steps 2-6, we define the following quantities:

$$
\begin{aligned}
n_0 &= \sum_{i=1}^{n}(1-z_i) \\
n_1 &= \sum_{i=1}^{n} z_i \\
\bar{y}_0 &= \frac{1}{n_0}\sum_{i=1}^{n}(1-z_i)y_i \\
\bar{y}_1 &= \frac{1}{n_1}\sum_{i=1}^{n} z_i y_i \\
SS_0 &= \sum_{i=1}^{n}(1-z_i)(y_i-\mu_0)^2 \\
SS_1 &= \sum_{i=1}^{n} z_i(y_i-\mu_1)^2
\end{aligned}
$$

2. Sampling the mixture probability α.

$$
\begin{aligned}
p(\alpha|\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y}) &\propto \left[\prod_{i=1}^{n} p(z_i|\alpha)\right] p(\alpha) \\
&= \prod_{i=1}^{n} \text{Bernoulli}(z_i|\alpha) \\
&= \prod_{i=1}^{n} \alpha^{z_i}(1-\alpha)^{1-z_i} \\
&= \alpha^{\sum_{i=1}^{n} z_i}(1-\alpha)^{n-\sum_{i=1}^{n} z_i} \\
&= \alpha^{n_1}(1-\alpha)^{n_0} \\
&= \text{Beta}(\alpha|n_1+1, n_0+1)
\end{aligned}
$$

3. Sampling the mixture mean μ_0.

$$
\begin{aligned}
p(\mu_0|\alpha, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y}) &\propto \left[\prod_{i=1}^{n} p(y_i|z_i=0, \mu_0, \sigma_0^2)\right] p(\mu_0, \sigma_0^2) \\
&\propto \prod_{\{i: z_i=0\}} N(y_i|\mu_0, \sigma_0^2) \\
&= \text{Normal}(\mu_0|\bar{y}_0, \sigma_0^2/n_0)
\end{aligned}
$$

4. Sampling the mixture mean μ_1.

$$\begin{aligned}
p(\mu_1|\alpha, \boldsymbol{\sigma}^2, \mathbf{z}, \mathbf{y}) &\propto \left[\prod_{i=1}^{n} p(y_i|z_i = 1, \mu_1, \sigma_1^2)\right] p(\mu_1, \sigma_1^2) \\
&\propto \prod_{\{i:z_i=1\}} N(y_i|\mu_1, \sigma_1^2) \\
&= \text{Normal}(\mu_1|\bar{y}_1, \sigma_1^2/n_1)
\end{aligned}$$

5. Sampling the mixture variance σ_0^2.

$$\begin{aligned}
p(\sigma_0^2|\alpha, \boldsymbol{\mu}, \mathbf{z}, \mathbf{y}) &\propto \left[\prod_{i=1}^{n} p(y_i|z_i = 0, \mu_0, \sigma_0^2)\right] p(\mu_0, \sigma_0^2) \\
&\propto \left[\prod_{\{i:z_i=0\}} N(y_i|\mu_0, \sigma_0^2)\right] (\sigma_0^2)^{-1} \\
&= \text{Inv-Gamma}\left(\frac{n_0}{2} + 1, \frac{SS_0}{2}\right)
\end{aligned}$$

6. Sampling the mixture variance σ_1^2.

$$\begin{aligned}
p(\sigma_1^2|\alpha, \boldsymbol{\mu}, \mathbf{z}, \mathbf{y}) &\propto \left[\prod_{i=1}^{n} p(y_i|z_i = 1, \mu_1, \sigma_1^2)\right] p(\mu_1, \sigma_1^2) \\
&\propto \left[\prod_{\{i:z_i=1\}} N(y_i|\mu_1, \sigma_1^2)\right] (\sigma_1^2)^{-1} \\
&= \text{Inv-Gamma}\left(\frac{n_1}{2} + 1, \frac{SS_1}{2}\right)
\end{aligned}$$

Finally, we note the correspondence between the Inv-χ^2 and Inv-Gamma distributions:

$$\text{Inv-}\chi^2(\nu, \hat{\sigma}^2) = \text{Inv-Gamma}\left(\frac{\nu}{2}, \frac{\nu\hat{\sigma}^2}{2}\right).$$