Project – Normal-Inverse Gamma Prior

Name:Liuqing Yang

March 5, 2018

Suppose we have p genes and for each gene, n subjects. The expression of gene $i, i = 1, ..., p$ and subject $j, j = 1, ...n$ is X_{ij}. Each gene belongs to a community k, denoted as class label $c_i = k$. We assume $X_{ij}|c_i = k \sim N(\mu_{ik}, \sigma_{ik}^2)$, where $(\mu_{ik}, \sigma_{ik}^2)$ has a Normal-Inverse Gamma prior with parameter $(\mu_{k0}, v_k, \alpha_k, \beta_k)$. That is:

$$f(x_{ij}|c_i = k, \mu_{ik}, \sigma_{ik}^2) = \frac{1}{\sqrt{2\pi}\sigma_{ik}} exp(-\frac{(x_{ij} - \mu_{ik})^2}{2\sigma_{ik}^2}),$$

$$f(\mu_{ik}, \sigma_{ik}^2|\mu_{k0}, v_k, \alpha_k, \beta_k) = \sqrt{\frac{v_k}{2\pi}} \frac{\beta_k^{\alpha_k}}{\Gamma(\alpha_k)} (\frac{1}{\sigma_{ik}^2})^{\alpha_k + \frac{1}{2}} exp\{-\frac{2\beta_k + v_k(\mu_{ik} - \mu_{ko})^2}{2\sigma_{ik}^2}\}$$

Let the expression of gene i as $X_i = (X_{i1}, ..., X_{in})^T$, then the conditional density of X_i is:

$$\begin{aligned} f(x_i|c_i = k, \mu_{ik}, \sigma_{ik}^2) &= \prod_{j=1}^{n} f(x_{ij}|c_i = k, \mu_{ik}, \sigma_{ik}^2) \\ &= \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_{ik}} exp(-\frac{(x_{ij} - \mu_{ik})^2}{2\sigma_{ik}^2}) \\ &= (2\pi)^{-\frac{n}{2}} (\frac{1}{\sigma_{ik}^2})^{\frac{n}{2}} exp\{-\frac{(\sum_{j=1}^{n} x_{ij}^2 - 2n\bar{x}_i\mu_{ik} + n\mu_{ik}^2)}{2\sigma_{ik}^2}\}, \quad \bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \end{aligned}$$

Then the marginal density of $X_i|c_i = k$ is:

$$f(x_i|c_i = k) = \frac{f(x_i|c_i = k, \mu_{ik}, \sigma_{ik}^2) f(\mu_{ik}, \sigma_{ik}^2)}{f(\mu_{ik}, \sigma_{ik}^2|x_i)},$$

since, in this case, Normal-Inverse Gamma is a conjugate prior, the posterior density of $\mu_{ik}, \sigma_{ik}^2|x_i$ is:

$$\begin{aligned} f(\mu_{ik}, \sigma_{ik}^2|x_i) &\propto f(x_i|c_i = k, \mu_{ik}, \sigma_{ik}^2) f(\mu_{ik}, \sigma_{ik}^2) \\ &= (2\pi)^{-\frac{n}{2}} (\frac{1}{\sigma_{ik}^2})^{\frac{n}{2}} exp\{-\frac{(\sum_{j=1}^{n} x_{ij}^2 - 2n\bar{x}_i\mu_{ik} + n\mu_{ik}^2)}{2\sigma_{ik}^2}\} \sqrt{\frac{v_k}{2\pi}} \frac{\beta_k^{\alpha_k}}{\Gamma(\alpha_k)} (\frac{1}{\sigma_{ik}^2})^{\alpha_k + \frac{1}{2}} exp\{-\frac{2\beta_k + v_k(\mu_{ik} - \mu_{ko})^2}{2\sigma_{ik}^2}\} \\ &\propto (\frac{1}{\sigma_{ik}^2})^{\alpha_k + \frac{n}{2} + \frac{1}{2}} exp\{-\frac{(v_k + n)\mu_{ik}^2 - 2(v_k\mu_{k0} + n\bar{x}_i)\mu_i k + v_k\mu_{k0}^2 + \sum_{j=1}^{n} x_{ij}^2 + 2\beta_k}{2\sigma_{ik}^2}\} \\ &= (\frac{1}{\sigma_{ik}^2})^{\alpha_k + \frac{n}{2} + \frac{1}{2}} exp\{-\frac{1}{2\sigma_{ik}^2}[(v_k + n)(\mu_{ik} - \frac{v_k\mu_{k0} + n\bar{x}_i}{v_k + n})^2 + v_k\mu_{k0}^2 + \sum_{j=1}^{n} x_{ij}^2 + 2\beta_k - \frac{1}{v_k + n}(v_k\mu_{k0} + n\bar{x}_i)^2]\}, \end{aligned}$$

where

$$\begin{aligned} v_k\mu_{k0}^2 + \sum_{j=1}^{n} x_{ij}^2 + 2\beta_k - \frac{1}{v_k + n}(v_k\mu_{k0} + n\bar{x}_i)^2 &= v_k\mu_{k0}^2 + \sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_i^2 + 2\beta_k - \frac{1}{v_k + n}(v_k\mu_{k0} + n\bar{x}_i)^2 + n\bar{x}_i^2 \\ &= 2\beta_k + \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 + \frac{nv_k}{v_k + n}(\mu_{ko} - \bar{x}_i)^2. \end{aligned}$$

Thus, $(\mu_{ik}, \sigma_{ik}^2|x_i)$ has a Normal-Inverse Gamma distribution with parameter:

$$(\frac{v_k\mu_{k0} + n\bar{x_i}}{v_k + n}, v_k + n, \alpha_k + \frac{n}{2}, \beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k + n}\frac{(\mu_{ko} - \bar{x_i})^2}{2}).$$

Then

$$f(x_i|c_i = k) = \frac{f(x_i|c_i = k, \mu_{ik}, \sigma_{ik}^2)f(\mu_{ik}, \sigma_{ik}^2)}{f(\mu_{ik}, \sigma_{ik}^2|x_i)}$$

$$\begin{aligned}
&= \frac{(2\pi)^{-\frac{n}{2}}(\frac{1}{\sigma_{ik}^2})^{\frac{n}{2}} exp\{-\frac{(\sum_n^{j=1} x_{ij}^2 - 2n\bar{x_i}\mu_{ik} + n\mu_{ik}^2)}{2\sigma_{ik}^2}\}\sqrt{\frac{v_k}{2\pi}}\frac{\beta_k^{\alpha_k}}{\Gamma(\alpha_k)}(\frac{1}{\sigma_{ik}^2})^{\alpha_k+\frac{1}{2}} exp\{-\frac{2\beta_k + v_k(\mu_{ik} - \mu_{ko})^2}{2\sigma_{ik}^2}\}}{\sqrt{\frac{v_k+n}{2\pi}}\frac{(\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko} - \bar{x_i})^2}{2})^{\alpha_k+\frac{n}{2}}}{\Gamma(\alpha_k + \frac{n}{2})}(\frac{1}{\sigma_{ik}^2})^{\alpha_k+\frac{n}{2}+\frac{1}{2}} exp\{-\frac{2\beta_k + \sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k+n}(\mu_{ko} - \bar{x_i})^2 + (v_k+n)(\mu_{ik} - \frac{v_k\mu_{k0}+n\bar{x_i}}{v_k+n})^2}{2\sigma_{ik}^2}\}} \\
&= \frac{(2\pi)^{-\frac{n}{2}}(\frac{1}{\sigma_{ik}^2})^{\frac{n}{2}}\sqrt{\frac{v_k}{2\pi}}\frac{\beta_k^{\alpha_k}}{\Gamma(\alpha_k)}(\frac{1}{\sigma_{ik}^2})^{\alpha_k+\frac{1}{2}}}{\sqrt{\frac{v_k+n}{2\pi}}\frac{(\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko} - \bar{x_i})^2}{2})^{\alpha_k+\frac{n}{2}}}{\Gamma(\alpha_k + \frac{n}{2})}(\frac{1}{\sigma_{ik}^2})^{\alpha_k+\frac{n}{2}+\frac{1}{2}}} \\
&= \frac{(2\pi)^{-\frac{n}{2}}\sqrt{\frac{v_k}{v_k+n}}\frac{\Gamma(\alpha_k+\frac{n}{2})}{\Gamma(\alpha_k)}\beta_k^{\alpha_k}}{(\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko} - \bar{x_i})^2}{2})^{\alpha_k+\frac{n}{2}}} := NE(\mu_{k0}, v_k, \alpha_k, \beta_k) = NE(x_i; \theta_k).
\end{aligned}$$

For E-M algorithm, suppose $p(c_i = k) = \pi_k, k = 1, ..., K, \pi = (\pi_1, ..., \pi_K)^T$ then,

$$\begin{aligned}
P(c_i = k|x_i) &= \frac{P(x_i|c_i = k)P(c_i = k)}{\sum_{k=1}^{K} P(x_i, c_i = k)} \\
&= \frac{NE(x_i; \theta_k)\pi_k}{\sum_{k=1}^{K} \pi_k NE(x_i; \theta_k)} \\
&:= T_{k,i}
\end{aligned}$$

$$\begin{aligned}
P(x, c|\theta) &= \prod_{i=1}^{p} P(x_i|c)P(c) \\
&= \prod_{i=1}^{p}[NE(x_i; \theta_k)]^{1(c_i=k)}\prod_{k=1}^{K}\pi_k^{n_k}, \quad n_k = \sum_{i=1}^{P} 1(c_i = k) \\
&= \prod_{i=1}^{p}\prod_{k=1}^{K}[NE(x_i; \theta_k)\pi_k]^{1(c_i=k)}.
\end{aligned}$$

E-Step:

$$\begin{aligned}
Q((\pi,\theta)|(\pi^{(t)},\theta^{(t)})) &= E_{C|X,\theta^{(t)}}[logL(\theta;X,C)] \\
&= \sum_{i=1}^{p}\sum_{k=1}^{K} P(c_i = k|X_i = x_i;\theta^{(t)})log(NE(x_i;\theta_k)\pi_k) \\
&= \sum_{i=1}^{p}\sum_{k=1}^{K} T_{k,i}[log\pi_k + log(NE(x_i;\theta_k))]
\end{aligned}$$

M-Step:

$$\pi_{(t+1)} = \arg\max_{\pi} \sum_{i=1}^{p}\sum_{k=1}^{K} T_{k,i}^{(t)} log\pi_k = \arg\max_{\pi} \sum_{k=1}^{K} log\pi_k \sum_{i=1}^{p} T_{k,i}^{(t)}$$

$$\Rightarrow \pi_k^{(t+1)} = \frac{\sum_{i=1}^{p} T_{k,i}^{(t)}}{\sum_{k=1}^{K}\sum_{i=1}^{p} T_{k,i}^{(t)}}$$

$$\begin{aligned}
\theta_k^{(t+1)} &= \arg\max_{\theta} \sum_{i=1}^{p} T_{k,i}^{(t)} logNE(x_i;\theta_k) \\
&= \arg\max_{\theta} \sum_{i=1}^{p} T_{k,i}^{(t)}[-\frac{n}{2}log2\pi + \frac{1}{2}log\frac{v_k}{v_k+n} + log\frac{\Gamma(_k+\frac{n}{2})}{\Gamma(\alpha_k)} + \alpha_k log\beta_k \\
&\quad - (\alpha_k + \frac{n}{2})log(\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko} - \bar{x_i})^2}{2})] \\
&:= \arg\max_{\theta} F(\theta_k)
\end{aligned}$$

$$\frac{\partial F(\theta_k)}{\partial \mu_{k0}} = \frac{\partial \sum_{i=1}^{p} T_{k,i}^{(t)} log(\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij}-\bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko}-\bar{x_i})^2}{2})]}{\partial \mu_k 0} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^{p} T_{k,i}^{(t)} \frac{nv_k}{v_k+n}(\mu_{ko} - \bar{x_i})}{\beta_k + \frac{1}{2}\sum_{j=1}^{n}(x_{ij}-\bar{x_i})^2 + \frac{nv_k}{v_k+n}\frac{(\mu_{ko}-\bar{x_i})^2}{2}} = 0$$

$$\Rightarrow \mu_{k0}^{(t+1)} = \frac{\sum_{i=1}^{p} T_{k,i}^{(t)}\bar{x_i}}{\sum_{i=1}^{p} T_{k,i}^{(t)}}$$