Bayesian hierarchical models - derivation notes

The Bayesian hierarchical model for the Kdrama rating application is:

- The data model:

$$Y_{ij} \mid \mu_j, \sigma \overset{i.i.d.}{\sim} \text{Normal}(\mu_j, \sigma), \;\; i = 1, \cdots, n_j, j = 1, \cdots J, \tag{1}$$

  where n_j is the number of observations in group j, and J is the number of groups.

- The prior distributions:

$$\begin{aligned} \mu_j \mid \mu, \tau &\overset{i.i.d.}{\sim} \text{Normal}(\mu, \tau) & (2)\\ \mu \mid \mu_0, \gamma_0 &\sim \text{Normal}(\mu_0, \gamma_0) & (3)\\ 1/\tau^2 \mid \alpha_\tau, \beta_\tau &\sim \text{Gamma}(\alpha_\tau, \beta_\tau) & (4)\\ 1/\sigma^2 \mid \alpha_\sigma, \beta_\sigma &\sim \text{Gamma}(\alpha_\sigma, \beta_\sigma) & (5) \end{aligned}$$

  We give fixed values for hyperparameters: $\mu_0, \gamma_0, \alpha_\tau, \beta_\tau, \alpha_\sigma, \beta_\sigma$, therefore we consider them as fixed and known. We also keep them in the conditions.

  The parameters in the model includes: μ_j's, σ, μ, and τ. Therefore, the likelihood function is:

- The likelihood:

$$\begin{aligned} L(\{\mu_j\}, \sigma, \mu, \tau) &= f(Y_{ij}, i = 1, \cdots, n_j, j = 1, \cdots, J \mid \{\mu_j\}, \sigma) \\ &= \prod_{j=1}^{J} \left(\prod_{i=1}^{n_j} f(Y_{ij} \mid \mu_j, \sigma) \right) \\ &= \prod_{j=1}^{J} \left(\prod_{i=1}^{n_j} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(y_{ij} - \mu_j)^2}{2\sigma^2} \right) \right) \end{aligned} \tag{6}$$

- The joint prior distribution:

$$\begin{aligned} \pi(\{\mu_j\}, \sigma, \mu, \tau \mid \mu_0, \gamma_0, \alpha_\tau, \beta_\tau, \alpha_\sigma, \beta_\sigma) &= \prod_{j=1}^{J} (\pi(\mu_j \mid \mu, \tau)) \pi(\mu \mid \mu_0, \gamma_0) \pi(\tau \mid \alpha_\tau, \beta_\tau) \pi(\sigma \mid \alpha_\sigma, \beta_\sigma) \\ &= \prod_{j=1}^{J} \left(\frac{1}{\sqrt{2\pi\tau^2}} \exp\left(- \frac{(\mu_j - \mu)^2}{2\tau^2} \right) \right) \\ &\quad \left(\frac{1}{\sqrt{2\pi\gamma_0^2}} \exp\left(- \frac{(\mu - \mu_0)^2}{2\gamma_0^2} \right) \right) \\ &\quad \frac{\beta_\tau^{\alpha_\tau}}{\Gamma(\alpha_\tau)} (1/\tau^2)^{\alpha_\tau - 1} \exp(-\beta_\tau (1/\tau^2)) \\ &\quad \frac{\beta_\sigma^{\alpha_\sigma}}{\Gamma(\alpha_\sigma)} (1/\sigma^2)^{\alpha_\sigma - 1} \exp(-\beta_\sigma (1/\sigma^2)) \end{aligned} \tag{7}$$

  (We work with $1/\tau^2$ and $1/\sigma^2$ instead of τ and σ.)

  With these pieces, we can express the joint posterior distribution using Bayes' theorem.

- The joint posterior distribution:

$$\begin{aligned}
\pi(\{\mu_j\},\sigma,\mu,\tau \mid \{y_{ij}\},\mu_0,\gamma_0,\alpha_\tau,\beta_\tau,\alpha_\sigma,\beta_\sigma) &\propto \pi(\{\mu_j\},\sigma,\mu,\tau \mid \mu_0,\gamma_0,\alpha_\tau,\beta_\tau,\alpha_\sigma,\beta_\sigma)L(\{\mu_j\},\sigma,\mu,\tau)\\
&\propto \prod_{j=1}^{J}\left(\prod_{i=1}^{n_j}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y_{ij}-\mu_j)^2}{2\sigma^2}\right)\right)\\
&\quad \prod_{j=1}^{J}\left(\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\mu_j-\mu)^2}{2\tau^2}\right)\right)\\
&\quad \left(\frac{1}{\sqrt{2\pi\gamma_0^2}}\exp\left(-\frac{(\mu-\mu_0)^2}{2\gamma_0^2}\right)\right)\\
&\quad \frac{\beta_\tau^{\alpha_\tau}}{\Gamma(\alpha_\tau)}(1/\tau^2)^{\alpha_\tau-1}\exp(-\beta_\tau(1/\tau^2))\\
&\quad \frac{\beta_\sigma^{\alpha_\sigma}}{\Gamma(\alpha_\sigma)}(1/\sigma^2)^{\alpha_\sigma-1}\exp(-\beta_\sigma(1/\sigma^2)) \qquad (8)
\end{aligned}$$

Next, to derive the full conditional posterior distribution for each parameter, we collect terms only related to the parameter we are working with in Equation (8). Other terms, either constants or terms related to other parameters, are considered given.

- The full conditional posterior distribution for μ:

$$\begin{aligned}
\pi(\mu \mid -) &\propto \exp\left(-\frac{\sum_{j=1}^{J}(\mu_j-\mu)^2}{2\tau^2}\right)\exp\left(-\frac{(\mu-\mu_0)^2}{2\gamma_0^2}\right)\\
&\propto \exp\left(-\frac{J\mu^2-2\mu\sum_{j=1}^{J}\mu_j}{2\tau^2}\right)\exp\left(-\frac{\mu^2-2\mu\mu_0}{2\gamma_0^2}\right)\\
&= \exp\left(-\frac{1}{2}\left(\left(\frac{J}{\tau^2}+\frac{1}{\gamma_0^2}\right)\mu^2-\left(\frac{2\sum_{j=1}^{J}\mu_j}{\tau^2}+\frac{2\mu_0}{\gamma_0^2}\right)\mu\right)\right) \qquad (9)
\end{aligned}$$

That is,

$$\mu \mid - \sim \text{Normal}\left(\frac{\sum_{j=1}^{J}\mu_j/\tau^2+\mu_0/\gamma_0^2}{J/\tau^2+1/\gamma_0^2},(J/\tau^2+1/\gamma_0^2)^{-1/2}\right). \qquad (10)$$

- The full conditional posterior distribution for $1/\tau^2$:

$$\begin{aligned}
\pi(1/\tau^2 \mid -) &\propto \prod_{j=1}^{J}(1/\tau^2)^{\frac{1}{2}}\exp\left(-\frac{\sum_{j=1}^{J}(\mu_j-\mu)^2}{2\tau^2}\right)(1/\tau^2)^{\alpha_\tau-1}\exp(-\beta_\tau(1/\tau^2))\\
&= (1/\tau^2)^{\frac{J}{2}}\exp\left(-\frac{\sum_{j=1}^{J}(\mu_j-\mu)^2}{2}(1/\tau^2)\right)(1/\tau^2)^{\alpha_\tau-1}\exp(-\beta_\tau(1/\tau^2))\\
&= (1/\tau^2)^{\frac{J}{2}+\alpha_\tau-1}\exp\left(-\left(\frac{\sum_{j=1}^{J}(\mu_j-\mu)^2}{2}+\beta_\tau\right)(1/\tau^2)\right) \qquad (11)
\end{aligned}$$

That is,

$$1/\tau^2 \mid - \sim \text{Gamma}\left(\alpha_\tau+\frac{J}{2},\beta_\tau+\frac{1}{2}\sum_{j=1}^{J}(\mu_j-\mu)^2\right). \qquad (12)$$

- The full conditional distribution for μ_j:

$$\begin{aligned}
\pi(\mu_j \mid -) &\propto \prod_{i=1}^{n_j} \exp\left(-\frac{(y_{ij}-\mu_j)^2}{2\sigma^2}\right) \exp\left(-\frac{(\mu_j-\mu)^2}{2\tau^2}\right) \\
&= \exp\left(-\frac{\sum_{i=1}^{n_j}(y_{ij}-\mu_j)^2}{2\sigma^2}\right) \exp\left(-\frac{(\mu_j-\mu)^2}{2\tau^2}\right) \\
&\propto \exp\left(-\frac{n_j\mu_j^2 - 2\mu_j\sum_{i=1}^{n_j} y_{ij}}{2\sigma^2}\right) \exp\left(-\frac{\mu_j^2 - 2\mu_j\mu}{2\tau^2}\right) \\
&= \exp\left(-\frac{1}{2}\left(\left(\frac{n_j}{\sigma^2}+\frac{1}{\tau^2}\right)\mu_j^2 - \left(\frac{2\sum_{i=1}^{n_j} y_{ij}}{\sigma^2}+\frac{2\mu}{\tau^2}\right)\mu_j\right)\right) \qquad (13)
\end{aligned}$$

That is,

$$\mu_j \mid - \sim \text{Normal}\left(\frac{\sum_{i=1}^{n_j} y_{ij}/\sigma^2 + \mu/\tau^2}{n_j/\sigma^2 + 1/\tau^2}, (n_j/\sigma^2+1/\tau^2)^{-1/2}\right). \qquad (14)$$

- The full conditional distribution for σ:

$$\begin{aligned}
\pi(1/\sigma^2 \mid -) &\propto \left(\prod_{j=1}^{J}\prod_{i=1}^{n_j}(1/\sigma^2)^{\frac{1}{2}}\right)\left(\prod_{j=1}^{J}\prod_{i=1}^{n_j}\exp\left(-\frac{(y_{ij}-\mu_j)^2}{2\sigma^2}\right)\right) \\
&\quad (1/\sigma^2)^{\alpha_\sigma-1}\exp(-\beta_\sigma(1/\sigma^2)) \\
&= (1/\sigma^2)^{\frac{\sum_{j=1}^{J} n_j}{2}} \exp\left(-(1/\sigma^2)\frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ij}-\mu_j)^2\right) \\
&\quad (1/\sigma^2)^{\alpha_\sigma-1}\exp(-\beta_\sigma(1/\sigma^2)) \\
&= (1/\sigma^2)^{\alpha_\sigma+\sum_{j=1}^{J} n_j/2-1} \exp\left(-(1/\sigma^2)\left(\beta_\sigma + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ij}-\mu_j)^2\right)\right) \qquad (15)
\end{aligned}$$

That is,

$$1/\sigma^2 \sim \text{Gamma}\left(\alpha_\sigma + \sum_{j=1}^{J}\frac{n_j}{2}, \beta_\sigma + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ij}-\mu_j)^2\right). \qquad (16)$$