VI : normal mean-precision

Sunday, 18 October 2020 5:58 pm

$$x_1, \dots x_n \sim N(\mu, \tfrac{1}{\tau})$$

prior : $p(\mu, \tau) = p(\mu|\tau)p(\tau)$

$$\mu|\tau \sim N(\mu_0, \tfrac{1}{\lambda_0 \tau})$$

$$\tau \sim \text{Gamma}(a_0, b_0)$$

$$p(\tau) = \frac{(b^0)^{a_0}}{\Gamma(a_0)} \tau^{a_0 - 1} e^{-b_0 \tau}$$

$$q(\mu, \tau) = q_\mu(\mu)\, q_\tau(\tau)$$

$$\log q_\mu^*(\mu) \propto E_\tau(\log p(\mu, \tau, x_1, \dots x_n)]$$

$$= E_\tau\left[\log \prod_{i=1}^{n} p(x_i|\tau, \mu)\, p(\mu|\tau) p(\tau)\right]$$

$$= E_\tau\left[\log\left(\sqrt{\tfrac{\tau}{2\pi}}\right)^n e^{-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2} + \log \sqrt{\tfrac{\lambda_0\tau}{2\pi}}\, e^{-\frac{\lambda_0\tau(\mu-\mu_0)^2}{2}}\right] + E_\tau(\log p(\tau))$$

$$\propto E_\tau\left[\frac{n}{2}\log\frac{\tau}{2\pi} - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1}{2}\log\frac{\lambda_0\tau}{2\pi} - \frac{\lambda_0\tau(\mu-\mu_0)^2}{2}\right]$$

$$\propto -\frac{E_\tau(\tau)}{2}\left[\sum_{i=1}^{n}(x_i-\mu)^2 + \lambda_0(\mu-\mu_0)^2\right]$$

$$\propto -\frac{E_\tau(\tau)}{2}\left[(\lambda_0+n)\mu^2 - 2\left(\lambda_0\mu_0 + \sum_{i=1}^{n} x_i\right)\mu\right]$$

$q_\mu^*(\mu)$ is a pdf of $N(\mu^*, \sigma^{2*})$,

$$\mu^* = \frac{\lambda_0\mu_0 + \sum_{i=1}^{n} x_i}{\lambda_0 + n}, \quad \sigma^{2*} = \frac{1}{(\lambda_0+n)\, E_\tau(\tau)}$$

$$\log q_\tau^*(\tau) \propto E_\mu(\log p(\mu, \tau, x_1, \dots x_n))$$

$$= E_\mu\left[\log \prod_{i=1}^{n} p(x_i|\mu, \tau)\, p(\mu|\tau) p(\tau)\right]$$

$$= E_\mu\left[\log\left(\sqrt{\tfrac{\tau}{2\pi}}\right)^n e^{-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2} + \log\sqrt{\tfrac{\lambda_0\tau}{2\pi}}\, e^{-\frac{\lambda_0\tau(\mu-\mu_0)^2}{2}}\right] + E_\mu(\log p(\tau))$$

$$= E_\mu\left[\frac{n}{2}\log\frac{\tau}{2\pi} - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1}{2}\log\frac{\lambda_0\tau}{2\pi} - \frac{\lambda_0\tau(\mu-\mu_0)^2}{2}\right] + \log p(\tau)$$

$$\propto \frac{n}{2}\log\tau - \frac{\tau}{2}\sum_{i=1}^{n} E_\mu[(x_i-\mu)^2] + \frac{1}{2}\log\tau - \frac{\lambda_0\tau}{2} E_\mu[(\mu-\mu_0)^2] + (a_0-1)\log\tau - b_0\tau$$

$$= \left(\frac{n+1}{2} + a_0 - 1\right)\log\tau - \tau\left[b_0 + \frac{1}{2}\sum_{i=1}^{n} E_\mu[(x_i-\mu)^2] + \frac{\lambda_0}{2} E_\mu[(\mu-\mu_0)^2]\right]$$

$q_\tau^*(\tau)$ is a pdf of $\text{Gamma}(a^*, b^*)$,

$$a^* = \frac{n+1}{2} + a_0, \quad b^* = b_0 + \frac{1}{2}\sum_{i=1}^{n} E_\mu[(x_i-\mu)^2] + \frac{\lambda_0}{2} E_\mu[(\mu-\mu_0)^2]$$

$$E_\tau(\tau) = \frac{a^*}{b^*}$$

$$E_\mu[(x_i-\mu)^2] = E_\mu(\mu^2) - 2x_i E_\mu(\mu) + x_i^2$$

$$= \sigma^{2*} + (\mu^*)^2 - 2x_i\mu^* + x_i^2 = A_i \cdots (1)$$

$$E_\mu[(\mu-\mu_0)^2] = E_\mu(\mu^2) - 2\mu_0 E_\mu(\mu) + \mu_0^2$$

$$= \sigma^{2*} + (\mu^*)^2 - 2\mu_0\mu^* + \mu_0^2 = B \cdots (2)$$

$$\text{ELBO}(q_\mu^*(\mu), q_\tau^*(\tau)) = \text{ELBO}(\mu^*, \sigma^{2*}, a^*, b^*)$$

$$= E_{\mu,\tau}\left(\log p(x_1 \dots x_n, \mu, \tau) - \log q_\mu^*(\mu) q_\tau^*(\tau)\right)$$

$$= E_{\mu,\tau}\left(\log \prod_{i=1}^{n} p(x_i|\mu,\tau) p(\mu|\tau) p(\tau)\right) - E_{\mu,\tau}(\log q_\mu^*(\mu)) - E_{\mu,\tau}(\log q_\tau^*(\tau))$$

$$E_{\mu,\tau}(\log q_\mu^*(\mu)) = E_{\mu,\tau}\left[-\frac{1}{2}\log\sigma^{2*} - \frac{1}{2}\log 2\pi - \frac{(\mu-\mu^*)^2}{2\sigma^{2*}}\right]$$

$$\propto -\frac{1}{2}\log\sigma^{2*} - \frac{1}{2\sigma^{2*}} E_\mu[(\mu-\mu^*)^2]$$

$$\propto -\frac{1}{2}\log\sigma^{2*} - \frac{\sigma^{2*}}{2\sigma^{2*}} \propto -\frac{1}{2}\log\sigma^{2*}$$

$$E_{\mu,\tau}(\log q_\tau^*(\tau)) = E_{\mu,\tau}\left(a^*\log b^* - \log\Gamma(a^*) + (a^*-1)\log\tau - b^*\tau\right)$$

$$= a^*\log b^* - \log\Gamma(a^*) + (a^*-1)E_\tau(\log\tau) - b^* E_\tau(\tau)$$

$$\boxed{\begin{aligned} E_\tau(\log\tau) &= -\log b^* + \psi(a^*) \\ \psi(x) &= \frac{\partial}{\partial x}\log\Gamma(x) \end{aligned}}$$

$$= a^*\log b^* - \log\Gamma(a^*) - (a^*-1)\log b^* + (a^*-1)\psi(a^*) - b^*\frac{a^*}{b^*}$$

$$= \log b^* - \log\Gamma(a^*) + (a^*-1)\psi(a^*) - a^*$$

$$E_{\mu,\tau}\left[\log \prod_{i=1}^{n} p(x_i|\mu,\tau)\, p(\mu|\tau) p(\tau)\right]$$

$$= \underbrace{E_{\mu,\tau}\left[\log\left(\sqrt{\tfrac{\tau}{2\pi}}\right)^n e^{-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2} + \log\sqrt{\tfrac{\lambda_0\tau}{2\pi}}\, e^{-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2}\right]}_{①} + \underbrace{E_\tau(\log p(\tau))}_{②}$$

①: $$E_{\mu,\tau}\left[\frac{n}{2}\log\tau - \frac{n}{2}\log 2\pi - \frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1}{2}\log\lambda_0 + \frac{1}{2}\log\tau - \frac{1}{2}\log 2\pi - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2\right]$$

$$\propto E_{\mu,\tau}\left[\frac{(n+1)}{2}\log\tau - \frac{\tau}{2}\left(\sum_{i=1}^{n}(x_i-\mu)^2 + \lambda_0(\mu-\mu_0)^2\right)\right]$$

$$= \frac{n+1}{2}E_\tau(\log\tau) - \frac{1}{2}E_\tau(\tau)\left[\sum_{i=1}^{n} E_\mu[(x_i-\mu)^2] + \lambda_0 E_\mu[(\mu-\mu_0)^2]\right]$$

$$= \frac{(n+1)}{2}\left(-\log b^* + \psi(a^*)\right) - \frac{1}{2}\frac{a^*}{b^*}\left[\sum_{i=1}^{n} A_i + \lambda_0 B\right]$$

A_i & B from (1) & (2)

②: $$E_\tau(\log p(\tau))$$

$$= E_\tau\left(a_0\log b_0 - \log\Gamma(a_0) + (a_0-1)\log\tau - b_0\tau\right)$$

$$\propto (a_0-1)E_\tau(\log\tau) - b_0 E_\tau(\tau)$$

$$= (a_0-1)\left(-\log b^* + \psi(a^*)\right) - b_0\frac{a^*}{b^*}$$