# VI : normal mean-precision

Sunday, 18 October 2020 5:58 pm

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

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

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

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

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

$q_i^*(\theta_i) \propto \exp\{E_{-i}[\log p(y|\theta) \cdot p(\theta)]\}$

$\Leftrightarrow$

$\log q_i^*(\theta_i) \propto E_{-i}[\log p(y|\theta) \cdot p(\theta)]$

---

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

---

$x_i$ is iid $\sim N(\mu, \frac{1}{\tau})$.

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

prior for $\tau$

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

"constant", doesn't contain $\mu$.

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

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

which part contain $\mu$ ?

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

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

$\Rightarrow -\frac{E_\tau(\tau)(\lambda_0+n)}{2}[\mu^2 - 2\frac{(\lambda_0\mu_0+\sum x_i)}{(\lambda_0+n)}\mu]$

$N(\mu, \sigma)$
$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

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

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

Not CHANGE for each iteration

change

---

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

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

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

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

cannot ignore since contain $\tau$

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

$\log p(\tau)$ & only keep $\tau$ part

$+ (a_0-1)\log\tau - b_0\tau$

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

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

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

No update ↑ keep updating

---

$E_\tau(\tau) = \frac{a^*}{b^*}$ $\int \tau q_\tau(\tau)\, d\tau$ since $\sim \text{Gamma}(a^*, b^*)$ $E_\tau(\tau) = \frac{a^*}{b^*}$

$E_\mu[(x_i-\mu)^2] = E_\mu(\mu^2) - 2x_i \boxed{E_\mu(\mu)} + x_i^2$ $\mu \sim N(\mu^*, \sigma^{2*})$.

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

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

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

---

Definition $\text{ELBO}(q(\theta)) = -\int_\theta q(\theta)[\log q(\theta) - \log \underbrace{p(\theta)p(y|\theta)}_{-p(\theta, y)}]d\theta$

$\text{ELBO}(q_\mu^*(\mu), q_\tau^*(\tau)) = \text{ELBO}(\mu^*, \sigma^{2*}, a^*, b^*) = \int_\theta q(\theta)[\log p(\theta, y) - \log q(\theta)]d\theta$

$= E_\theta[\log p(\theta, y) - \log q(\theta)]$

with respect to $q(\theta)$

$= E_{\mu,\tau}(\log p(x_1 \cdots x_n, \mu, \tau) - \log \underbrace{\boxed{q_\mu^*(\mu) q_\tau^*(\tau)}}_{q(\theta)})$ mean field variational family

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

---

$\mu \sim N(\mu^*, \sigma^{2*})$.

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

$\propto -\frac{1}{2}\log\sigma^{2*} - \frac{1}{2\sigma^{2*}} E_\mu[(\mu-\mu^*)^2]$ (since no $\tau$ inside) $\sim N(\mu^*, \sigma^{2*})$.

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

---

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

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

for $\tau \sim$ Gamma

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

$\psi(x) = \frac{\partial}{\partial x}\log\Gamma(x)$

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

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

---

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

should be $E_{\mu,\tau}$ but ignore $\mu$ since ( ) doesn't contain $\mu$

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

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

$+ \frac{1}{2}\log\lambda_0 + \frac{1}{2}\log\tau - \frac{1}{2}\log 2\pi - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2]]$

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

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

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

$A_i$ & $B$ from (1) & (2)

$E_{\mu,\tau}(f(\mu)\, g(\tau))$
$= E_\tau[E_\mu[f(\mu)\, g(\tau)]]$
$= E_\tau[g(\tau)E_\mu(f(\mu))]$
$= E_\mu(f(\mu))\, E_\tau(g(\tau))$

②: $E_\tau(\log p(\tau))$ log of Gamma density

$= E_\tau(\boxed{\underbrace{a_0\log b_0 - \log\Gamma(a_0)}_{\text{"constant" with respect to } \tau} + (a_0-1)\log\tau - b_0\tau})$

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

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