Cosine basis

Daeyoung Lim*
Department of Statistics
Korea University

March 29, 2016

1 Normal

1.1 Model specifications

$$
\begin{aligned}
y_i &= w_i^\top \beta + f\left(x_i\right) + \epsilon_i, \qquad \epsilon_i \sim \mathcal{N}\left(0, \sigma^2\right) \\
\theta_j | \sigma, \tau, \gamma &\sim \mathcal{N}\left(0, \sigma^2 \tau^2 \exp\left[-j\gamma\right]\right) \\
\tau^2 &\sim \mathrm{IG}\left(\frac{r_{0,\tau}}{2}, \frac{s_{0,\tau}}{2}\right) \\
\sigma^2 &\sim \mathrm{IG}\left(\frac{r_{0,\sigma}}{2}, \frac{s_{0,\sigma}}{2}\right) \\
\beta &\sim \mathcal{N}\left(\mu_\beta^0, \Sigma_\beta^0\right) \\
\gamma &\sim \mathrm{Exp}\left(\omega_0\right) \\
|\psi| &= \gamma, \quad \psi \sim \mathrm{DE}\left(0, \omega_0\right) \\
\varphi_j\left(x\right) &= \sqrt{2}\cos\left(\pi j x\right)
\end{aligned}
$$

Joint density:

$$
p\left(y, \Theta\right) = \mathcal{N}\left(y \middle| W\beta + f_J, \sigma^2 I_n\right) \left\{ \prod_{j=1}^{J} \mathcal{N}\left(\theta_j \middle| 0, \sigma^2 \tau^2 \exp\left[-j\left|\psi\right|\right]\right) \right\} \mathrm{IG}\left(\tau^2 \middle| \frac{r_{0,\tau}}{2}, \frac{s_{0,\tau}}{2}\right) \mathrm{IG}\left(\sigma^2 \middle| \frac{r_{0,\sigma}}{2}, \frac{s_{0,\sigma}}{2}\right) \mathcal{N}\left(\beta \middle| \mu_\beta^0, \Sigma_\beta^0\right) \mathrm{DE}\left(\psi | 0, \omega_0\right)
$$

We will use the joint density to derive the LB and updating algorithm. The variational distributions are

$$
\begin{aligned}
q_1\left(\beta\right) &= \mathcal{N}\left(\mu_\beta^q, \Sigma_\beta^q\right) \\
q_2\left(\theta_J\right) &= \mathcal{N}\left(\mu_\theta^q, \Sigma_\theta^q\right) \\
q_3\left(\sigma^2\right) &= \mathrm{IG}\left(\frac{r_{q,\sigma}}{2}, \frac{s_{q,\sigma}}{2}\right) \\
q_4\left(\tau^2\right) &= \mathrm{IG}\left(\frac{r_{q,\tau}}{2}, \frac{s_{q,\tau}}{2}\right) \\
q_5\left(\psi\right) &= \mathcal{N}\left(\mu_\psi^q, {\sigma_\psi^2}^q\right) \quad \text{(NCVMP)}.
\end{aligned}
$$

*Prof. Taeryon Choi

1.2 Lower bound

1.2.1 LB: $\mathsf{E}\left[\ln p\left(y|\Theta\right)\right]$

$$\begin{aligned}\mathsf{E}\left[\ln p\left(y|\Theta\right)\right] &= -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2}\mathsf{E}\left[\left(y - W\beta - \varphi_J\theta\right)^\top\left(y - W\beta - \varphi_J\theta\right)\right] \\ &= -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2}\left(y - W\mu_\beta^q - \varphi_J\mu_\theta^q\right)^\top\left(y - W\mu_\beta^q - \varphi_J\mu_\theta^q\right) - \frac{1}{2}\left(\mathrm{Tr}\left(W^\top W\Sigma_\beta^q\right) + \mathrm{Tr}\left(\varphi_J^\top\varphi_J\Sigma_\theta^q\right)\right)\end{aligned}$$

1.3 LB: $\mathsf{E}\left[\ln p\left(\theta_j|\sigma,\tau,\psi\right)\right]$

$$\sum_{j=1}^{J}\mathsf{E}\left[\ln p\left(\theta_j|\sigma,\tau,\psi\right)\right] = \sum_{j=1}^{J}\mathsf{E}\left[-\frac{1}{2}\ln\left(2\pi\right) + \ln\frac{1}{\sigma} + \ln\frac{1}{\tau} + \frac{j}{2}\left|\psi\right| - \frac{\theta_j^2 e^{j|\psi|}}{2\sigma^2\tau^2}\right]$$

Let's note the following fact: if $X\sim\mathcal{N}\left(\mu,\sigma^2\right)$, then $|X|\sim$ folded-Normal $\left(\mu,\sigma^2\right)$. Then,

$$\begin{aligned}\mathsf{E}\left|X\right| &= \sigma\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\} + \mu\left(1 - 2\Phi\left(\frac{-\mu}{\sigma}\right)\right) \\ &= \sigma\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\} - \mu\,\mathrm{erf}\left(\frac{-\mu}{\sqrt{2\sigma^2}}\right) \\ \mathsf{E}e^{t|X|} &= \exp\left\{\frac{\sigma^2t^2}{2} + \mu t\right\}\left[1 - \Phi\left(-\frac{\mu}{\sigma} - \sigma t\right)\right] + \exp\left\{\frac{\sigma^2t^2}{2} - \mu t\right\}\left[1 - \Phi\left(\frac{\mu}{\sigma} - \sigma t\right)\right].\end{aligned}$$

2 Probit

2.1 Model specifications

$$\begin{aligned}\Pr\left(y_i = 1|f,\beta\right) &= \Phi\left(w_i^\top\beta + f\left(x_i\right)\right) \\ y_i^* &= w_i^\top\beta + f\left(x_i\right) + \epsilon_i, \qquad \epsilon_i\sim\mathcal{N}\left(0,1\right) \\ y_i &= \begin{cases}1, & \text{if } y_i^*\geq 0 \\ 0, & \text{if } y_i^* < 0\end{cases} \\ \theta_j|\sigma,\tau,\gamma &\sim\mathcal{N}\left(0,\sigma^2\tau^2\exp\left[-j\gamma\right]\right) \\ \tau^2 &\sim\mathrm{IG}\left(\frac{r_{0,\tau}}{2},\frac{s_{0,\tau}}{2}\right) \\ \sigma^2 &\sim\mathrm{IG}\left(\frac{r_{0,\sigma}}{2},\frac{s_{0,\sigma}}{2}\right) \\ \beta &\sim\mathcal{N}\left(\mu_\beta^0,\Sigma_\beta^0\right) \\ \gamma &\sim\mathrm{Exp}\left(\omega_0\right) \\ |\psi| &= \gamma, \quad \psi\sim\mathrm{DE}\left(0,\omega_0\right) \\ \varphi_j\left(x\right) &= \sqrt{2}\cos\left(\pi jx\right)\end{aligned}$$

Joint density:

$$\begin{aligned}p\left(y,y^*,\Theta\right) = C&\left\{\prod_{j=1}^{J}\mathcal{N}\left(\theta_j|0,\sigma^2\tau^2\exp\left[-j\left|\psi\right|\right]\right)\right\}\mathrm{IG}\left(\tau^2\middle|\frac{r_{0,\tau}}{2},\frac{s_{0,\tau}}{2}\right)\mathrm{IG}\left(\sigma^2\middle|\frac{r_{0,\sigma}}{2},\frac{s_{0,\sigma}}{2}\right)\mathcal{N}\left(\beta|\mu_\beta^0,\Sigma_\beta^0\right) \\ &\mathrm{DE}\left(\psi|0,\omega_0\right)\left\{\prod_{i=1}^{n}\left(1\left[y_i^*\geq 0\right]1\left[y_i = 1\right] + 1\left[y_i^* < 0\right]1\left[y_i = 0\right]\right)\right\}\phi\left(y_i^* - w_i^\top\beta - \varphi_i^\top\theta_J\right)\end{aligned}$$

where C is the normalizing constant. The variational distributions are

$$\begin{aligned}
q_1(\beta) &= \mathcal{N}\left(\mu^q_\beta, \Sigma^q_\beta\right)\\
q_2(\theta_J) &= \mathcal{N}\left(\mu^q_\theta, \Sigma^q_\theta\right)\\
q_3\left(\sigma^2\right) &= \mathrm{IG}\left(\frac{r_{q,\sigma}}{2}, \frac{s_{q,\sigma}}{2}\right)\\
q_4\left(\tau^2\right) &= \mathrm{IG}\left(\frac{r_{q,\tau}}{2}, \frac{s_{q,\tau}}{2}\right)\\
q_5(\psi) &= \mathcal{N}\left(\mu^q_\psi, \sigma^{2\,q}_\psi\right) \quad \text{(NCVMP)}\\
q_6(y^*) &= \mathcal{TN}\left(\mu^q_{y^*}, I_n, 0\right)
\end{aligned}$$

2.2 Lower bound

2.2.1 LB: $\mathsf{E}\left[\ln p\left(y^*|\mathbf{rest}\right)\right] + \mathsf{H}\left[y^*\right]$

$$\begin{aligned}
\mathsf{E}\left[\ln p\left(y^*|\text{rest}\right)\right] + \mathsf{H}\left[y^*\right] &= \sum_{i=1}^n \mathsf{E}\left[\ln\phi\left(y_i^* - w_i^\top\beta - \varphi_i^\top\theta_J\right) - \ln\phi\left(y_i^* - w_i^\top\mu^q_\beta - \varphi_i^\top\mu^q_\theta\right)\right]\\
&\quad + \sum_{i=1}^n \ln\left(\left\{\Phi\left(w_i^\top\mu^q_\beta + \varphi_i^\top\mu^q_\theta\right)\right\}^{y_i}\left\{1 - \Phi\left(w_i^\top\mu^q_\beta + \varphi_i^\top\mu^q_\theta\right)\right\}^{1-y_i}\right)\\
&= -\frac{1}{2}\left(\mathrm{Tr}\left(W^\top W\Sigma^q_\beta\right) + \mathrm{Tr}\left(\varphi_J^\top\varphi_J\Sigma^q_\theta\right)\right)\\
&\quad + \sum_{i=1}^n \ln\left(\left\{\Phi\left(w_i^\top\mu^q_\beta + \varphi_i^\top\mu^q_\theta\right)\right\}^{y_i}\left\{1 - \Phi\left(w_i^\top\mu^q_\beta + \varphi_i^\top\mu^q_\theta\right)\right\}^{1-y_i}\right)
\end{aligned}$$

2.3 LB: $\mathsf{E}\left[\ln p\left(\theta_j|\sigma,\tau,\psi\right)\right]$

Let's note the following fact: if $X \sim \mathcal{N}\left(\mu,\sigma^2\right)$, then $|X| \sim \text{folded-Normal}\left(\mu,\sigma^2\right)$. Then,

$$\begin{aligned}
\mathsf{E}\,|X| &= \sigma\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\} + \mu\left(1 - 2\Phi\left(\frac{-\mu}{\sigma}\right)\right)\\
&= \sigma\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{\mu^2}{2\sigma^2}\right\} - \mu\,\mathrm{erf}\left(\frac{-\mu}{\sqrt{2\sigma^2}}\right)\\
\mathsf{E}e^{t|X|} &= \exp\left\{\frac{\sigma^2t^2}{2} + \mu t\right\}\left[1 - \Phi\left(-\frac{\mu}{\sigma} - \sigma t\right)\right] + \exp\left\{\frac{\sigma^2t^2}{2} - \mu t\right\}\left[1 - \Phi\left(\frac{\mu}{\sigma} - \sigma t\right)\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{j=1}^J \mathsf{E}\left[\ln p\left(\theta_j|\sigma,\tau,\psi\right)\right] + \mathsf{H}\left[\theta_J\right] &= \sum_{j=1}^J \mathsf{E}\left[-\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\frac{1}{\sigma^2} + \frac{1}{2}\ln\frac{1}{\tau^2} + \frac{j}{2}|\psi| - \frac{\theta_j^2 e^{j|\psi|}}{2\sigma^2\tau^2}\right] + \mathsf{H}\left[\theta_J\right]\\
&= -\frac{J}{2}\left\{\ln(2\pi) - \left(\mathrm{di}\left(\frac{r_{q,\sigma}}{2}\right) - \ln\left(\frac{s_{q,\sigma}}{2}\right)\right) - \left(\mathrm{di}\left(\frac{r_{q,\tau}}{2}\right) - \ln\left(\frac{s_{q,\tau}}{2}\right)\right)\right\}\\
&\quad + \frac{J(J+1)}{4}\left\{\sigma^q_\psi\sqrt{\frac{2}{\pi}}\exp\left(-\frac{\mu^{q\,2}_\psi}{2\sigma^{q\,2}_\psi}\right) + \mu^q_\psi\left(1 - 2\Phi\left(\frac{-\mu^q_\psi}{\sigma^q_\psi}\right)\right)\right\}\\
&\quad - \frac{1}{2}\frac{r_{q,\sigma}}{s_{q,\sigma}}\frac{r_{q,\tau}}{s_{q,\tau}}\left(\mathrm{Tr}\left(\Sigma^q_\theta\right) + \mu^{q\top}_\theta\mu^q_\theta\right)\sum_{j=1}^J Q_j\left(\mu^q_\psi, \sigma^{q\,2}_\psi\right) + \frac{J}{2}\left(1 + \ln(2\pi)\right) + \frac{1}{2}\ln\left|\Sigma^q_\theta\right|
\end{aligned}$$

where

$$
\begin{aligned}
Q_j\left(\mu_\psi^q, {\sigma_\psi^q}^2\right) &= \mathsf{E} e^{j|\psi|} \\
&= \exp\left\{\frac{{\sigma_\psi^q}^2 j^2}{2} + \mu_\psi^q j\right\}\left[1 - \Phi\left(-\frac{\mu_\psi^q}{\sigma_\psi^q} - \sigma_\psi^q j\right)\right] + \exp\left\{\frac{{\sigma_\psi^q}^2 j^2}{2} - \mu_\psi^q j\right\}\left[1 - \Phi\left(\frac{\mu_\psi^q}{\sigma_\psi^q} - \sigma_\psi^q j\right)\right].
\end{aligned}
$$