Cosine basis

Daeyoung Lim*
Department of Statistics
Korea University

June 10, 2016

1 Normal

1.1 Model specifications

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

Joint density:

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

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

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

*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,

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

2 Probit: No Restriction

2.1 Model specifications

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

Joint density:

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

where C is the normalizing constant. The variational distributions are

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

2.2 Lower bound

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

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

2.2.2 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 $|X| \sim$ folded-Normal $\left(\mu, \sigma^2\right)$. Then,

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

Therefore,

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

where

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

2.2.3 LB: $\mathsf{E}\left[\ln p\left(\tau^2\right)\right] + \mathsf{H}\left[\tau^2\right]$

$$
\begin{aligned}
\mathsf{E}\left[\ln p\left(\tau^2\right)\right] + \mathsf{H}\left[\tau^2\right] = &\frac{r_{0,\tau}}{2} \ln \frac{s_{0,\tau}}{2} - \ln \Gamma\left(\frac{r_{0,\tau}}{2}\right) + \left(\frac{r_{0,\tau}}{2} + 1\right)\left\{\mathrm{di}\left(\frac{r_{q,\tau}}{2}\right) - \ln\left(\frac{s_{q,\tau}}{2}\right)\right\} - \frac{s_{0,\tau}}{2}\frac{r_{q,\tau}}{s_{q,\tau}} \\
&+ \frac{r_{q,\tau}}{2} + \ln \frac{s_{q,\tau}}{2} + \ln \Gamma\left(\frac{r_{q,\tau}}{2}\right) - \left(1 + \frac{r_{q,\tau}}{2}\right)\mathrm{di}\left(\frac{r_{q,\tau}}{2}\right)
\end{aligned}
$$

2.2.4 LB: $\mathsf{E}\left[\ln p\left(\sigma^2\right)\right] + \mathsf{H}\left[\sigma^2\right]$

$$
\begin{aligned}
\mathsf{E}\left[\ln p\left(\sigma^2\right)\right] + \mathsf{H}\left[\sigma^2\right] = &\frac{r_{0,\sigma}}{2} \ln \frac{s_{0,\sigma}}{2} - \ln \Gamma\left(\frac{r_{0,\sigma}}{2}\right) + \left(\frac{r_{0,\sigma}}{2} + 1\right)\left\{\mathrm{di}\left(\frac{r_{q,\sigma}}{2}\right) - \ln\left(\frac{s_{q,\sigma}}{2}\right)\right\} - \frac{s_{0,\sigma}}{2}\frac{r_{q,\sigma}}{s_{q,\sigma}} \\
&+ \frac{r_{q,\sigma}}{2} + \ln \frac{s_{q,\sigma}}{2} + \ln \Gamma\left(\frac{r_{q,\sigma}}{2}\right) - \left(1 + \frac{r_{q,\sigma}}{2}\right)\mathrm{di}\left(\frac{r_{q,\sigma}}{2}\right)
\end{aligned}
$$

2.2.5 LB: $\mathsf{E}\left[\ln p\left(\beta\right)\right] + \mathsf{H}\left[\beta\right]$

$$
\begin{aligned}
\mathsf{E}\left[\ln p\left(\beta\right)\right] + \mathsf{H}\left[\beta\right] = &\frac{p+1}{2} + \frac{1}{2}\left(\mathrm{di}\left(\frac{r_{q,\sigma}}{2}\right) - \ln\left(\frac{s_{q,\sigma}}{2}\right)\right) + \frac{1}{2}\ln\left|{\Sigma_\beta^0}^{-1}\Sigma_\beta^q\right| \\
&- \frac{1}{2}\frac{r_{q,\sigma}}{s_{q,\sigma}}\left\{\mathrm{Tr}\left({\Sigma_\beta^0}^{-1}\Sigma_\beta^q\right) + \left(\mu_\beta^q - \mu_\beta^0\right)^\top {\Sigma_\beta^0}^{-1}\left(\mu_\beta^q - \mu_\beta^0\right)\right\}
\end{aligned}
$$

2.2.6 LB: $\mathsf{E}\left[\ln p\left(\psi\right)\right] + \mathsf{H}\left[\psi\right]$

$$
\begin{aligned}
\mathsf{E}\left[\ln p\left(\psi\right)\right] + \mathsf{H}\left[\psi\right] = &\ln \frac{\omega_0}{2} - \omega_0\left\{\sigma_\psi^q \sqrt{\frac{2}{\pi}}\exp\left(-\frac{{\mu_\psi^q}^2}{2{\sigma_\psi^q}^2}\right) + \mu_\psi^q\left(1 - 2\Phi\left(-\frac{\mu_\psi^q}{\sigma_\psi^q}\right)\right)\right\} \\
&+ \frac{1}{2}\ln\left(2\pi{\sigma_\psi^q}^2\right) - \frac{1}{2}
\end{aligned}
$$

2.3 Update

2.3.1 θ_j

$$
\begin{aligned}
\Sigma_\theta^q &\leftarrow \left(\varphi_J^\top \varphi_J + \frac{r_{q,\sigma}}{s_{q,\sigma}}\frac{r_{q,\tau}}{s_{q,\tau}}\mathrm{Dg}\left(Q_{1:J}\right)\right)^{-1} \\
\mu_\theta^q &\leftarrow \Sigma_\theta^q \varphi_J^\top\left(\mu_{y^*}^q - W\mu_\beta^q\right)
\end{aligned}
$$

2.3.2 τ^2

$$
\begin{aligned}
r_{q,\tau} &\leftarrow r_{0,\tau} + J \\
s_{q,\tau} &\leftarrow s_{0,\tau} + \frac{r_{q,\sigma}}{s_{q,\sigma}}\sum_{j=1}^{J}\left(\Sigma_{\theta,jj}^q + {\mu_{\theta,j}^q}^2\right)Q_j
\end{aligned}
$$

2.3.3 σ^2

$$
\begin{aligned}
r_{q,\sigma} &\leftarrow r_{0,\sigma} + J + p + 1 \\
s_{q,\sigma} &\leftarrow s_{0,\sigma} + \frac{r_{q,\tau}}{s_{q,\tau}} \sum_{j=1}^{J} \left(\Sigma_{\theta,jj}^{q} + {\mu_{\theta,j}^{q}}^{2} \right) Q_j + \mathrm{Tr}\left({\Sigma_{\beta}^{0}}^{-1} \Sigma_{\beta}^{q} \right) + \left(\mu_{\beta}^{q} - \mu_{\beta}^{0} \right)^{\top} {\Sigma_{\beta}^{0}}^{-1} \left(\mu_{\beta}^{q} - \mu_{\beta}^{0} \right)
\end{aligned}
$$

2.3.4 β

$$
\begin{aligned}
\Sigma_{\beta}^{q} &\leftarrow \frac{s_{q,\sigma}}{r_{q,\sigma}} \left(W^{\top} W + {\Sigma_{\beta}^{0}}^{-1} \right)^{-1} \\
\mu_{\beta}^{q} &\leftarrow \frac{r_{q,\sigma}}{s_{q,\sigma}} \Sigma_{\beta}^{q} \left({\Sigma_{\beta}^{0}}^{-1} \mu_{\beta}^{0} + W^{\top} \left(\mu_{y^*}^{q} - \varphi_J \mu_{\theta}^{q} \right) \right)
\end{aligned}
$$

2.3.5 ψ

$$
\begin{aligned}
\frac{\partial S_1}{\partial \mu_{\psi}^{q}} &= -\omega_0 \left\{ -\frac{1}{\sigma_{\psi}^{q}} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{{\mu_{\psi}^{q}}^{2}}{2{\sigma_{\psi}^{q}}^{2}} \right) + 1 - 2\Phi\left(-\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} \right) + 2\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} \phi\left(-\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} \right) \right\} \\
\frac{\partial S_1}{\partial {\sigma_{\psi}^{q}}^{2}} &= -\omega_0 \left\{ \left(\frac{1}{\sqrt{2\pi}\sigma_{\psi}^{q}} + \frac{{\mu_{\psi}^{q}}^{2}}{\sqrt{\pi}\left({\sigma_{\psi}^{q}}^{2}\right)^{3/2}} \right) \exp\left(-\frac{{\mu_{\psi}^{q}}^{2}}{{\sigma_{\psi}^{q}}^{2}} \right) - \phi\left(-\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} \right) \frac{{\mu_{\psi}^{q}}^{2}}{\left({\sigma_{\psi}^{q}}^{2}\right)^{3/2}} \right\} \\
\frac{\partial Q_j}{\partial \mu_{\psi}^{q}} &= j \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\} + \frac{1}{\sigma_{\psi}^{q}} \exp\left(\frac{{\sigma_{\psi}^{q}}^{2} j^2}{2} + \mu_{\psi}^{q} j \right) \phi\left(-\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} - \sigma_{\psi}^{q} j \right) \\
&\quad - j \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\} - \frac{1}{\sigma_{\psi}^{q}} \exp\left(\frac{{\sigma_{\psi}^{q}}^{2} j^2}{2} - \mu_{\psi}^{q} j \right) \phi\left(\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} - \sigma_{\psi}^{q} j \right) \\
\frac{\partial Q_j}{\partial {\sigma_{\psi}^{q}}^{2}} &= j \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\} + \left(\frac{1}{2\sigma_{\psi}^{q}} j - \frac{\mu_{\psi}^{q}}{2\left({\sigma_{\psi}^{q}}^{2}\right)^{3/2}} \right) \phi\left(-\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} - \sigma_{\psi}^{q} j \right) \exp\left(\frac{{\sigma_{\psi}^{q}}^{2} j^2}{2} + \mu_{\psi}^{q} j \right) \\
&\quad + j \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\} + \left(\frac{1}{2\sigma_{\psi}^{q}} j + \frac{\mu_{\psi}^{q}}{2\left({\sigma_{\psi}^{q}}^{2}\right)^{3/2}} \right) \phi\left(\frac{\mu_{\psi}^{q}}{\sigma_{\psi}^{q}} - \sigma_{\psi}^{q} j \right) \exp\left(\frac{{\sigma_{\psi}^{q}}^{2} j^2}{2} - \mu_{\psi}^{q} j \right) \\
\frac{\partial S_2}{\partial \mu_{\psi}^{q}} &= -\frac{1}{2} \frac{r_{q,\sigma}}{s_{q,\sigma}} \frac{r_{q,\tau}}{s_{q,\tau}} \left(\mathrm{Tr}\left(\Sigma_{\theta}^{q} \right) + {\mu_{\theta}^{q}}^{\top} \mu_{\theta}^{q} \right) \sum_{j=1}^{J} \frac{\partial Q_j}{\partial \mu_{\psi}^{q}} - \frac{J\left(J+1\right)}{4\omega_0} \frac{\partial S_1}{\partial \mu_{\psi}^{q}} \\
\frac{\partial S_2}{\partial {\sigma_{\psi}^{q}}^{2}} &= -\frac{1}{2} \frac{r_{q,\sigma}}{s_{q,\sigma}} \frac{r_{q,\tau}}{s_{q,\tau}} \left(\mathrm{Tr}\left(\Sigma_{\theta}^{q} \right) + {\mu_{\theta}^{q}}^{\top} \mu_{\theta}^{q} \right) \sum_{j=1}^{J} \frac{\partial Q_j}{\partial {\sigma_{\psi}^{q}}^{2}} - \frac{J\left(J+1\right)}{4\omega_0} \frac{\partial S_1}{\partial {\sigma_{\psi}^{q}}^{2}} \\
{\sigma_{\psi}^{q}}^{2} &\leftarrow -\frac{1}{2} \left\{ \frac{\partial S_1}{\partial {\sigma_{\psi}^{q}}^{2}} + \frac{\partial S_2}{\partial {\sigma_{\psi}^{q}}^{2}} \right\}^{-1} \\
\mu_{\psi}^{q} &\leftarrow \mu_{\psi}^{q} + {\sigma_{\psi}^{q}}^{2} \left\{ \frac{\partial S_1}{\partial \mu_{\psi}^{q}} + \frac{\partial S_2}{\partial \mu_{\psi}^{q}} \right\}
\end{aligned}
$$

3 Probit: Monotone Shape Restriction

3.1 Model

$$
\begin{aligned}
\mathrm{P}\left(y_i=1|\beta,\theta_J\right) &= \Phi\left(w_i^\top\beta+\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)\\
y_i^* &= w_i^\top\beta+\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J+\epsilon_i, \quad \epsilon_i\sim\mathcal{N}\left(0,1\right)\\
y_i &= \begin{cases}1, & \text{if } y_i^*\geq 0\\ 0, & \text{if } y_i^*<0\end{cases}\\
\theta_j|\sigma,\tau,\gamma &\sim \mathcal{N}\left(0,\sigma\tau^2\exp\left[-j\gamma\right]\right), \quad \text{for } j\geq 2\\
\theta_0|\sigma &\sim \mathcal{N}\left(0,\sigma\sigma_0^2\right)\\
\tau^2 &\sim \mathrm{IG}\left(\frac{r_{0,\tau}}{2},\frac{s_{0,\tau}}{2}\right)\\
\sigma^2 &\sim \mathrm{IG}\left(\frac{r_{0,\sigma}}{2},\frac{s_{0,\sigma}}{2}\right)\\
\beta &\sim \mathcal{N}\left(\mu_\beta^0,\sigma^2\Sigma_\beta^0\right) \qquad (p\times 1)\\
\gamma &\sim \mathrm{Exp}\left(\omega_0\right)\\
|\psi| &= \gamma, \quad \psi\sim\mathrm{DE}\left(0,\omega_0\right)\\
\varphi_{0,0}^a\left(x\right) &= x-0.5\\
\varphi_{0,j}^a\left(x\right) &= \varphi_{j,0}^a\left(x\right)=\frac{\sqrt{2}}{\pi j}\sin\left(\pi jx\right)-\frac{\sqrt{2}}{\left(\pi j\right)^2}\left[1-\cos\left(\pi j\right)\right] \text{ for } j\geq 1,\\
\varphi_{j,j}^a\left(x\right) &= \frac{\sin\left(2\pi jx\right)}{2\pi j}+x-0.5 \text{ for } j\geq 1,\\
\varphi_{j,k}^a\left(x\right) &= \frac{\sin\left[\pi\left(j+k\right)x\right]}{\pi\left(j+k\right)}+\frac{\sin\left[\pi\left(j-k\right)x\right]}{\pi\left(j-k\right)}\\
&\quad -\frac{1-\cos\left[\pi\left(j+k\right)\right]}{\left[\pi\left(j+k\right)\right]^2}-\frac{1-\cos\left[\pi\left(j-k\right)\right]}{\left[\pi\left(j-k\right)\right]^2}\\
&\quad \text{for } j\neq k \text{ and } j,k\geq 1.
\end{aligned}
$$

Joint density:

$$
\begin{aligned}
p\left(y,y^*,\Theta\right) \propto\ & \mathcal{N}\left(\theta_0|0,\sigma\sigma_0^2\right)\left\{\prod_{j=1}^J\mathcal{N}\left(\theta_j|0,\sigma\tau^2\exp\left[-j\left|\psi\right|\right]\right)\right\}\mathrm{IG}\left(\tau^2\middle|\frac{r_{0,\tau}}{2},\frac{s_{0,\tau}}{2}\right)\mathrm{IG}\left(\sigma^2\middle|\frac{r_{0,\sigma}}{2},\frac{s_{0,\sigma}}{2}\right)\mathcal{N}\left(\beta|\mu_\beta^0,\sigma^2\Sigma_\beta^0\right)\\
&\times\mathrm{DE}\left(\psi|0,\omega_0\right)\left\{\prod_{i=1}^n\left(1\left[y_i^*\geq 0\right]1\left[y_i=1\right]+1\left[y_i^*<0\right]1\left[y_i=0\right]\right)\phi\left(y_i^*-w_i^\top\beta-\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)\right\}
\end{aligned}
$$

3.2 Update

3.2.1 y_i^*

$$
\begin{aligned}
\ln q\left(y_i^*\right) &\propto -\frac{1}{2}\,\mathrm{E}_{-y_i^*}\left(y_i^{*2}-2y_i^*\left(w_i^\top\beta+\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)+\left(w_i^\top\beta+\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)^2\right)\\
&\propto -\frac{1}{2}\left(y_i^{*2}-2y_i^*\left(w_i^\top\mu_\beta^q+\delta\left(\mathrm{Tr}\left(\varphi_J^a\left(x_i\right)\Sigma_\theta^q\right)+\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta^q\right)\right)\right.\\
&\quad+\mathrm{Tr}\left(w_iw_i^\top\Sigma_\beta^q\right)+\mu_\beta^{q\top}w_iw_i^\top\mu_\beta^q+2\delta w_i^\top\mu_\beta^q\left(\mathrm{Tr}\left(\varphi_J^a\left(x_i\right)\Sigma_\theta^q\right)+\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta^q\right)\\
&\quad\left.+2\delta^2\,\mathrm{Tr}\left(\varphi_J^a\left(x_i\right)\Sigma_\theta^q\right)^2+4\delta^2\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\Sigma_\theta^q\varphi_J^a\left(x_i\right)\mu_\theta^q+\delta^2\left(\mathrm{Tr}\left(\varphi_J^a\left(x_i\right)\Sigma_\theta^q\right)+\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta^q\right)^2\right)\\
&\propto -\frac{1}{2}\left(y_i^*-w_i^\top\mu_\beta^q-\delta\left(\mathrm{Tr}\left(\varphi_J\left(x_i\right)\Sigma_\theta^q\right)+\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta\right)\right)^2
\end{aligned}
$$

Therefore,

$$
q\left(y_i^*\right)\sim\mathcal{TN}\left(\mu_{y_i^*},1\right)
$$

truncated at zero where

$$
\mu_{y_i^*}=w_i^\top\mu_\beta^q+\delta\,\mathrm{Tr}\left(\varphi_J\left(x_i\right)\Sigma_\theta^q\right)+\delta\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta.
$$

To distinguish the mean of the normal distribution and the mean of the truncated normal distribution, the latter will be denoted $\mu_{y_i^*}^q$ where

$$
\mathrm{E}\left(y_i^*\right)=\mu_{y_i^*}^q=\mu_{y_i^*}+\frac{\phi\left(\mu_{y_i^*}\right)}{\Phi\left(\mu_{y_i^*}\right)^{y_i}\left(\Phi\left(\mu_{y_i^*}\right)-1\right)^{1-y_i}}.
$$

3.2.2 θ

$$
\begin{aligned}
\ln q\left(\theta\right) &\propto \mathrm{E}_{-\theta}\left[-\frac{1}{2}\sum_{j=1}^{J}\frac{e^{j|\psi|}\theta_j^2}{\sigma^2\tau^2}-\frac{1}{2}\sum_{i=1}^{n}\left(y_i^*-w_i^\top\beta-\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)^2\right]\\
&\propto -\frac{1}{2}\,\mathrm{E}\left(\frac{1}{\sigma_2}\right)\mathrm{E}\left(\frac{1}{\tau^2}\right)\theta_J^\top\,\mathrm{Dg}\left(Q_{1:J}\right)\theta_J\\
&\quad-\frac{1}{2}\sum_{i=1}^{n}\mathrm{E}\left(\left(w_i^\top\beta+\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)^2-2y_i^*\left(w_i^\top\beta+\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)\right)
\end{aligned}
$$

Seems like the authors of the original paper used NVBMP.

3.2.3 β

$$
\begin{aligned}
\ln q\left(\beta\right) &\propto \mathrm{E}_{-\beta}\left(-\frac{1}{2\sigma^2}\left(\beta-\mu_\beta^0\right)^\top\Sigma_\beta^{0\,-1}\left(\beta-\mu_\beta^0\right)-\frac{1}{2}\sum_{i=1}^{n}\left(y_i^*-w_i^\top\beta-\delta\theta_J^\top\varphi_J^a\left(x_i\right)\theta_J\right)^2\right)\\
&\propto -\frac{1}{2}\left(\beta^\top\left(\mathrm{E}\left(\frac{1}{\sigma^2}\right)\Sigma_\beta^{0\,-1}+W^\top W\right)\beta\right.\\
&\quad\left.-2\left(\mathrm{E}\left(\frac{1}{\sigma^2}\right)\beta^\top\Sigma_\beta^{0\,-1}\mu_\beta^0+\beta^\top W^\top\mu_{y^*}^q-\beta^\top\delta\sum_{i=1}^{n}w_i\left(\mathrm{Tr}\left(\varphi_J^a\left(x_i\right)\Sigma_\theta^q\right)+\mu_\theta^{q\top}\varphi_J^a\left(x_i\right)\mu_\theta^q\right)\right)\right)
\end{aligned}
$$

Therefore, $q(\beta) = \mathcal{N}\left(\mu_\beta^q, \Sigma_\beta^q\right)$ where

$$\Sigma_\beta^q = \left(\mathrm{E}\left(\frac{1}{\sigma^2}\right)\Sigma_\beta^{0\,-1} + W^\top W\right)^{-1}$$
$$\mu_\beta^q = \Sigma_\beta^q\left(\mathrm{E}\left(\frac{1}{\sigma^2}\right)\Sigma_\beta^{0\,-1}\mu_\beta^0 + W^\top \mu_{y^*}^q - \delta\sum_{i=1}^n w_i\left(\mathrm{Tr}\left(\varphi_J^a(x_i)\,\Sigma_\theta^q\right) + \mu_\theta^{q\top}\varphi_J^a(x_i)\,\mu_\theta^q\right)\right)$$

3.2.4 τ^2

$$r_{q,\tau} = r_{0,\tau} + J$$
$$s_{q,\tau} = s_{0,\tau} + \mathrm{E}\left(\frac{1}{\sigma}\right)\mathrm{Tr}\left(\left(\Sigma_\theta^{q*} + \mu_\theta^{q*}\mu_\theta^{q*\top}\right)\mathrm{Dg}\left(Q_{1:J}\right)\right)$$

3.2.5 σ^2

$$\ln q\left(\sigma^2\right) \propto -\frac{J}{2}\ln\sigma - \frac{1}{2}$$