GP Probit

Daeyoung Lim*
Department of Statistics
Korea University

April 1, 2016

1 GP Probit with sinusoidal basis

$$
\begin{aligned}
\alpha|\sigma &\sim \mathcal{N}\left(0, \frac{\sigma^2}{m} I_{2m}\right) \\
\lambda &\sim \mathcal{N}\left(\mu_\lambda, \Sigma_\lambda\right) \\
\sigma &\sim \text{half-Cauchy}\left(A_\sigma\right) \\
\theta &= \left(\alpha, \lambda, \sigma\right)
\end{aligned}
$$

$$
\pi\left(y, y^*, \theta\right) = C\pi\left(\alpha|\sigma\right)\pi\left(\lambda\right)\pi\left(\sigma\right)\prod_{i=1}^{n}\left\{1\left[y_i^* \geq 0\right]1\left[y_i = 1\right] + 1\left[y_i^* < 0\right]1\left[y_i = 0\right]\right\}\phi\left(y_i^* - Z_i^\top\alpha\right)
$$

1.1 Mean-field update

1.1.1 $q\left(y^*\right)$

$$
\mu_{y^*}^q \leftarrow Z\mu_\alpha^q + \frac{\phi\left(Z\mu_\alpha^q\right)}{\left\{\Phi\left(Z\mu_\alpha^q\right)\right\}^y\left\{\Phi\left(Z\mu_\alpha^q\right) - 1_n\right\}^{1_n - y}}
$$

1.1.2 $q\left(\alpha\right)$

$$
\begin{aligned}
\Sigma_\alpha^q &\leftarrow m\,\mathrm{E}\left[\frac{1}{\sigma^2}\right]I_{2m} + \mathrm{E}\left[Z^\top Z\right] \\
\mu_\alpha^q &\leftarrow \Sigma_\alpha^q\,\mathrm{E}\left[Z\right]^\top\mu_{y^*}^q
\end{aligned}
$$

1.1.3 $q\left(\sigma\right)$

$$
\begin{aligned}
q\left(\sigma\right) &\propto \frac{\exp\left(-C_\sigma/\sigma^2\right)}{\sigma^{2m}\left(A_\sigma^2 + \sigma^2\right)} \\
C_\sigma &\leftarrow \frac{m}{2}\left(\mathrm{Tr}\left(\Sigma_\alpha^q\right) + {\mu_\alpha^q}^\top\mu_\alpha^q\right)
\end{aligned}
$$

*Prof. Taeryon Choi

1.2 Lower bound

1.2.1 LB: $\mathrm{E}\left[\ln p\left(y, y^{*} \mid \theta\right)\right]+\mathrm{H}\left[y^{*}\right]$

$$
\begin{aligned}
\mathrm{E}\left[\ln p\left(y, y^{*} \mid \theta\right)\right]+\mathrm{H}\left[y^{*}\right]= & -\frac{1}{2}\left\{\operatorname{Tr}\left(\mathrm{E}\left[Z^{\top} Z\right] \Sigma_{\alpha}^{q}\right)+\mu_{\alpha}^{q \top}\left(\mathrm{E}\left[Z^{\top} Z\right]-\mathrm{E}[Z]^{\top} \mathrm{E}[Z]\right) \mu_{\alpha}^{q}\right\} \\
& +\sum_{i=1}^{n} \ln \left(\left\{\Phi\left(\mathrm{E}\left[Z_{i}\right]^{\top} \mu_{\alpha}^{q}\right)\right\}^{y_{i}}\left\{1-\Phi\left(\mathrm{E}\left[Z_{i}\right]^{\top} \mu_{\alpha}^{q}\right)\right\}^{1-y_{i}}\right)
\end{aligned}
$$

1.2.2 LB: $\mathrm{E}[\ln \pi(\alpha \mid \sigma)]+\mathrm{H}[\alpha]$

$$
\mathrm{E}[\ln \pi(\alpha \mid \sigma)]+\mathrm{H}[\alpha]=-m(2 \mathrm{E}[\ln \sigma]-\ln m)-\frac{m}{2} \mathrm{E}\left[\frac{1}{\sigma^{2}}\right]\left(\operatorname{Tr}\left(\Sigma_{\alpha}^{q}\right)+\mu_{\alpha}^{q \top} \mu_{\alpha}^{q}\right)+m+\frac{1}{2} \ln \left|\Sigma_{\alpha}^{q}\right|
$$

1.2.3 LB: $\mathrm{E}[\ln \pi(\lambda)]+\mathrm{H}[\lambda]$

$$
\mathrm{E}[\ln \pi(\lambda)]+\mathrm{H}[\lambda]=\frac{1}{2} \ln \left|\Sigma_{\lambda}^{-1} \Sigma_{\lambda}^{q}\right|-\frac{1}{2}\left\{\operatorname{Tr}\left(\Sigma_{\lambda}^{-1} \Sigma_{\lambda}^{q}\right)+\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)^{\top} \Sigma_{\lambda}^{-1}\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)\right\}+\frac{d}{2}
$$

1.2.4 LB: $\mathrm{E}[\ln \pi(\sigma)]+\mathrm{H}[\sigma]$

$$
\mathrm{E}[\ln \pi(\sigma)]+\mathrm{H}[\sigma]=\ln \left(\frac{2 A_{\sigma}}{\pi}\right)+C_{\sigma} \frac{\mathcal{H}\left(2 m, C_{\sigma}, A_{\sigma}^{2}\right)}{\mathcal{H}\left(2 m-2, C_{\sigma}, A_{\sigma}^{2}\right)}+2 m \mathrm{E}[\ln \sigma]+\ln \mathcal{H}\left(2 m-2, C_{\sigma}, A_{\sigma}^{2}\right)
$$

$$
\begin{aligned}
\mathcal{L}= & -\frac{1}{2}\left\{\operatorname{Tr}\left(\mathrm{E}\left[Z^{\top} Z\right] \Sigma_{\alpha}^{q}\right)+\mu_{\alpha}^{q \top}\left(\mathrm{E}\left[Z^{\top} Z\right]-\mathrm{E}[Z]^{\top} \mathrm{E}[Z]\right) \mu_{\alpha}^{q}\right\} \\
& +\sum_{i=1}^{n} \ln \left(\left\{\Phi\left(\mathrm{E}\left[Z_{i}\right]^{\top} \mu_{\alpha}^{q}\right)\right\}^{y_{i}}\left\{1-\Phi\left(\mathrm{E}\left[Z_{i}\right]^{\top} \mu_{\alpha}^{q}\right)\right\}^{1-y_{i}}\right) \\
& +m \ln m-\frac{m}{2} \mathrm{E}\left[\frac{1}{\sigma^{2}}\right]\left(\operatorname{Tr}\left(\Sigma_{\alpha}^{q}\right)+\mu_{\alpha}^{q \top} \mu_{\alpha}^{q}\right)+m+\frac{1}{2} \ln \left|\Sigma_{\alpha}^{q}\right| \\
& +\frac{1}{2} \ln \left|\Sigma_{\lambda}^{-1} \Sigma_{\lambda}^{q}\right|-\frac{1}{2}\left\{\operatorname{Tr}\left(\Sigma_{\lambda}^{-1} \Sigma_{\lambda}^{q}\right)+\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)^{\top} \Sigma_{\lambda}^{-1}\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)\right\}+\frac{d}{2} \\
& +\ln \left(\frac{2 A_{\sigma}}{\pi}\right)+C_{\sigma} \frac{\mathcal{H}\left(2 m, C_{\sigma}, A_{\sigma}^{2}\right)}{\mathcal{H}\left(2 m-2, C_{\sigma}, A_{\sigma}^{2}\right)}+\ln \mathcal{H}\left(2 m-2, C_{\sigma}, A_{\sigma}^{2}\right)
\end{aligned}
$$

1.3 NCVMP: λ update

$$
S_{a}=-\frac{1}{2}\left\{\operatorname{Tr}\left(\Sigma_{\lambda}^{-1} \Sigma_{\lambda}^{q}\right)+\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)^{\top} \Sigma_{\lambda}^{-1}\left(\mu_{\lambda}^{q}-\mu_{\lambda}\right)\right\}+y^{* \top} \mathrm{E}[Z] \mu_{\alpha}^{q}-\frac{1}{2}\left(\mathrm{E}\left[Z^{\top} Z\right]\left(\Sigma_{\alpha}^{q}+\mu_{\alpha}^{q} \mu_{\alpha}^{q \top}\right)\right)
$$