1 Stochastic Variational Inference in Gaussian Regression

For the sprectral decomposition of the Gaussian process, we use the following expansion:

$$f(x) \approx \sum_{r=1}^{m} \left[a_r \cos\left\{ (s_r \odot x)' \lambda \right\} + b_r \sin\left\{ (s_r \odot x)' \lambda \right\} \right] \tag{1}$$

where $\lambda = (\lambda_1, \ldots, \lambda_d)'$ ad $\odot$ denots the Hadamard product.

- $\alpha = (a_1, \ldots, a_m, b_1, \ldots, b_m)'$
- $y = (y_1, \ldots, y_n)'$
- $Z = (Z_1, \ldots, Z_n)'$
- $Z_i = (\cos((s_1 \odot x_i)'\lambda), \ldots, \cos((s_m \odot x_i)'\lambda), \sin((s_1 \odot x_i)'\lambda), \ldots, \sin((s_m \odot x_i)'\lambda)$
- $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$

Then,

$$y = Z\alpha + \epsilon, \quad \epsilon \sim \mathcal{N}\left(0, \gamma^2 I_n\right) \tag{2}$$

with the following priors

- $\alpha \sim \mathcal{N}\left(0, \frac{\sigma^2}{m} I_{2m}\right)$
- $\lambda \sim \mathcal{N}\left(\mu_\lambda^0, \Sigma_\lambda^0\right)$
- $\sigma \sim \mathrm{HF}\left(A_\sigma\right)$
- $\gamma \sim \mathrm{HF}\left(A_\gamma\right)$

where the density of half-Cauchy distribution is as follows:

$$\pi(\sigma) = \frac{2A_\sigma}{\pi\left(A_\sigma^2 + \sigma^2\right)}. \tag{3}$$

The fixed variational posteriors are

- $q(\alpha) = \mathcal{N}\left(\mu_\alpha^q, \Sigma_\alpha^q\right)$
- $q(\lambda) = \mathcal{N}\left(\mu_\lambda^q, \Sigma_\lambda^q\right)$

- $q(\sigma) = b_\sigma^{a_\sigma} / \Gamma\left(a_\sigma\right) \left(\sigma^2\right)^{-(a_\sigma+1)} e^{-b_\sigma/\sigma}$
- $q(\gamma) = b_\gamma^{a_\gamma} / \Gamma\left(a_\gamma\right) \left(\gamma^2\right)^{-(a_\gamma+1)} e^{-b_\gamma/\gamma}$

1.1 Lower bound

Let $h(\theta) = p(y \mid \theta)\pi(\theta)$ where $\theta = (\alpha, \lambda, \sigma, \gamma)$.

$$\begin{aligned}
\log h(\theta) = &-\frac{n}{2}\log\gamma^2 - \frac{1}{2\gamma^2}(y - Z\alpha)'(y - Z\alpha) && (4)\\
&- m\log\frac{\sigma^2}{m} - \frac{m}{2\sigma^2}\alpha'\alpha && (5)\\
&+ \log(2A_\sigma) - \log\left(\pi\left(A_\sigma^2 + \sigma^2\right)\right) + \log\left(2A_\gamma\right) - \log\left(\pi\left(A_\gamma^2 + \gamma^2\right)\right) && (6)\\
&+ \frac{1}{2}\log\left|\Sigma_\lambda^{0^{-1}}\right| - \frac{1}{2}\left(\lambda - \mu_\lambda^0\right)'\Sigma_\lambda^{0^{-1}}\left(\lambda - \mu_\lambda^0\right) && (7)
\end{aligned}$$

The variational posterior is denoted by $q_\lambda(\theta)$.

$$\begin{aligned}
\log q_\lambda(\theta) = &-m\log|\Sigma_\alpha^q| - \frac{1}{2}\left(\alpha - \mu_\alpha^q\right)'\Sigma_\alpha^{q^{-1}}\left(\alpha - \mu_\alpha^q\right) && (8)\\
&- \frac{d}{2}\log\left|\Sigma_\lambda^q\right| - \frac{1}{2}\left(\lambda - \mu_\lambda^q\right)'\Sigma_\lambda^{q^{-1}}\left(\lambda - \mu_\lambda^q\right) && (9)\\
&+ a_\sigma\log b_\sigma - \log\Gamma\left(a_\sigma\right) - \left(a_\sigma + 1\right)\log\sigma^2 - \frac{b_\sigma}{\sigma^2} && (10)\\
&+ a_\gamma\log b_\gamma - \log\Gamma\left(a_\gamma\right) - \left(a_\gamma + 1\right)\log\gamma^2 - \frac{b_\gamma}{\gamma^2} && (11)
\end{aligned}$$