BAYES HW 3

Vayser Kirill

November 2021

1 Problem 1

1.1 1 point

Using Jensen inequality:

$$\begin{aligned}
\mathcal{L}_k &\leq \log \mathbb{E}_{z_1,\cdots,z_k \sim q_\phi(z|x)} \frac{1}{k} \sum_{i=1}^{k} \frac{p_\theta(x \mid z_i) p(z_i)}{q_\phi(z_i \mid x)} = \\
&= \log \mathbb{E}_{z \sim q_\phi(z|x)} \frac{p_\theta(x \mid z) p(z)}{q_\phi(z \mid x)} = \log \int \frac{p_\theta(x \mid z) p(z)}{q_\phi(z \mid x)} q_\phi(z \mid x) dz = \\
&= \log \int p_\theta(x \mid z) p(z) dz = \log p(x)
\end{aligned}$$

1.2 2 point

Let w_i^z be w_i from problem statement and w_i^ε be result of reparametrization:

$$w_i^\varepsilon = \frac{p_\theta(x \mid g_\phi(\varepsilon_i)) p(g_\phi(\varepsilon_i))}{q_\phi(g_\phi(\varepsilon_i) \mid x)}$$

$$\nabla_{\theta,\phi} \mathcal{L}_1 = \nabla_{\theta,\phi} \mathbb{E}_{z \sim q_\phi(z|x)} \log w^z = \nabla_{\theta,\phi} \mathbb{E}_{\varepsilon \sim p(\varepsilon)} \log w^\varepsilon = \mathbb{E}_{\varepsilon \sim p(\varepsilon)} \nabla_{\theta,\phi} \log w^\varepsilon$$

$$\begin{aligned}
\nabla_{\theta,\phi} \mathcal{L}_k &= \nabla_{\theta,\phi} \mathbb{E}_{z_1,\cdots,z_k \sim q_\phi(z|x)} \log \frac{1}{k} \sum_{i=1}^{k} w_i^z = \nabla_{\theta,\phi} \mathbb{E}_{\varepsilon_1,\cdots,\varepsilon_k \sim p(\varepsilon)} \log \frac{1}{k} \sum_{i=1}^{k} w_i^\varepsilon = \\
&= \mathbb{E}_{\varepsilon_1,\cdots,\varepsilon_k \sim p(\varepsilon)} \nabla_{\theta,\phi} \log \frac{1}{k} \sum_{i=1}^{k} w_i^\varepsilon = \mathbb{E}_{\varepsilon_1,\cdots,\varepsilon_k \sim p(\varepsilon)} \frac{\sum_{i=1}^{k} \nabla_{\theta,\phi} w_i^\varepsilon}{\sum_{i=1}^{k} w_i^\varepsilon} = \\
&= \mathbb{E}_{\varepsilon_1,\cdots,\varepsilon_k \sim p(\varepsilon)} \frac{\sum_{i=1}^{k} w_i^\varepsilon \nabla_{\theta,\phi} \log w_i^\varepsilon}{\sum_{i=1}^{k} w_i^\varepsilon}
\end{aligned}$$

2 Problem 2

$$\begin{aligned}\mathbb{E}_{p_e(x)}\mathcal{L}(x) &= \sum_{n=1}^{N}\int \log\frac{p_\theta(x_n \mid z)p(z)}{q_\phi(z \mid x_n)}q_\phi(z \mid x_n)dz = A + \sum_{n=1}^{N}\int \log p(z)q_\phi(z \mid x_n)dz = \\ &= A + N\int \log p(z)\overline{q}_\phi(z \mid x_n)dz = A - N\int \log\frac{\overline{q}_\phi(z \mid x_n)}{p(z)}\overline{q}_\phi(z \mid x_n)dz + \\ &+ N\int \log \overline{q}_\phi(z \mid x_n)\overline{q}_\phi(z \mid x_n)dz = B - KL(\overline{q}\|p) \Rightarrow p(z)^* = \overline{q}(z),\end{aligned}$$

where $\overline{q} = \frac{1}{N}\sum_{n=1}^{N} q_\phi(z \mid x_n)$