Understanding the Final Objective Equation of CorrSH

Zijing Ou[1]

Theorem 0.1. *In corrSH [1], the final objective function is defined as*

$$\mathcal{L}_k = \mathbb{E}_{q_\phi(s|x)}\left[\log \frac{p_\theta(x|s)p(s)}{\mathrm{e}^{-E_\phi(s)}}\right] - \mathbb{E}_{q_\phi(r^{1\ldots k}|x)}\left[\mathbb{E}_{h_k(s)}\left[\log \frac{h_k(s)}{\mathrm{e}^{-E_\phi(s)}}\right]\right], \tag{1}$$

where $h_k(s) = \frac{1}{k}\sum_i q_\phi(s|r^{(i)})$, and $q_\phi(s|x) = \int q_\phi(s|r)q_\phi(r|x)\mathrm{d}r$. In this script, we show that the (1) can be estimated as the following unbiased Monte Carlo sampling technique

$$\mathcal{L}_k \approx \log p_\theta(x|s) - \sum_{i=1}^{d} \log 2 - \operatorname{logsumexp}\left(\sum_{j=1}^{d}\left(\tilde{s}_j r_j^{(i)} - \operatorname{softplus}(r_j^{(i)})\right) - \log k\right),$$

$$\text{where } s \sim q_\phi(s|x),\ \tilde{s} \sim h_k(\tilde{s}) = \frac{1}{k}\sum_{i=1}^{k} q_\phi(\tilde{s}|r^{(i)}),\ r^{(1\ldots k)} \sim q_\phi(r^{(1\ldots k)}|x).$$

Proof. To show this, we first simplify (1) by canceling out the exponential term. Denote $A(s) := \log \mathrm{e}^{-E_\phi(s)}$, we have

$$\begin{aligned}
& - \mathbb{E}_{q_\phi(s|x)}\left[\log \mathrm{e}^{E_\phi(s)}\right] + \mathbb{E}_{q_\phi(r^{1\ldots k}|x)}\left[\mathbb{E}_{h_k(s)}\left[-\log \mathrm{e}^{-E_\phi(s)}\right]\right] \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \mathbb{E}_{q_\phi(r^{1\ldots k}|x)}\left[\mathbb{E}_{h_k(s)}\left[-A(s)\right]\right] \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \int q_\phi(r^{(1\ldots k)}|x)\left[\sum_s A(s)\left(\frac{1}{k}\sum_{i=1}^{k} q_\phi(s|r^{(i)})\right)\right]\mathrm{d}r^{(1)}\ldots\mathrm{d}r^{(k)} \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \sum_s A(s)\left(\frac{1}{k}\sum_{i=1}^{k}\int q_\phi(r^{(i)}|x)q_\phi(s|r^{(i)})\mathrm{d}r^{(i)}\right) \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \sum_s A(s)\left(\frac{1}{k}\sum_{i=1}^{k} q_\phi(s|x)\right) \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \sum_s A(s)q_\phi(s|x) \\
&= -\mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] + \mathbb{E}_{q_\phi(s|x)}\left[A(s)\right] = 0.
\end{aligned} \tag{2}$$

Thereby, we have

$$\mathcal{L}_k = \mathbb{E}_{q_\phi(s|x)}\left[\log p_\theta(x|s)p(s)\right] - \mathbb{E}_{q_\phi(r^{1\ldots k}|x)}\left[\mathbb{E}_{h_k(s)}\left[\log h_k(s)\right]\right], \tag{3}$$

which can be estimated unbiasedly by using Monte Carlo sampling

$$\begin{gathered}
\mathcal{L}_k \approx \log p_\theta(x|s) + \log p(s) - \log h_k(\tilde{s}), \\
\text{where } s \sim q_\phi(s|x),\ \tilde{s} \sim h_k(\tilde{s}) = \frac{1}{k}\sum_{i=1}^{k} q_\phi(\tilde{s}|r^{(i)}),\ r^{(1\ldots k)} \sim q_\phi(r^{(1\ldots k)}|x).
\end{gathered} \tag{4}$$

[1] ouzj@mail2.sysu.edu.cn

Here, we consider one Monte Carlo sample solely. One can apply multiple samples to reduce variance further. Next, we show how to further express (4) in closed form.

First, since we use Bernoulli prior, the second term in (4) can be rewrited as

$$\begin{aligned}\log p(s) &= \log \mathrm{Ber}(s; p = 0.5) \\ &= \log \prod_{i=1}^{d} 0.5^{s_i}(1-0.5)^{1-s_i} \\ &= -\sum_{i=1}^{d} \log 2. \end{aligned} \tag{5}$$

For the third term, we have

$$\begin{aligned}\log h_k(\tilde{s}) &= \log \frac{1}{k} \sum_{i=1}^{k} q_\phi(\tilde{s}|r^{(i)}) \\ &= \underbrace{\log \sum_{i=1}^{k} \exp}_{\text{logsumexp}} \left(\log q_\phi(\tilde{s}|r^{(i)}) - \log k\right). \end{aligned} \tag{6}$$

Note that $\log q_\phi(\tilde{s}|r^{(i)})$ can be expressed in closed form. First, we denote $q_\phi(\tilde{s}|r^{(i)})$ as

$$\begin{aligned}q_\phi(\tilde{s}|r^{(i)}) &= \mathrm{Ber}(\tilde{s}; p = \sigma(r^{(i)})) \\ &= \prod_{j=1}^{d} \sigma(r_j^{(i)})^{\tilde{s}_j}(1-\sigma(r_j^{(i)}))^{1-\tilde{s}_j}, \end{aligned} \tag{7}$$

where $\sigma(\cdot)$ denotes the sigmoid function. Then, we have

$$\begin{aligned}\log q_\phi(\tilde{s}|r^{(i)}) &= \sum_{j=1}^{d} \tilde{s}_j \log(\sigma(r_j^{(i)})) + (1-\tilde{s}_j)\log(1-\sigma(r_j^{(i)})) \\ &= \sum_{j=1}^{d} \tilde{s}_j \log \frac{1}{1+\mathrm{e}^{-r_j^{(i)}}} + (1-\tilde{s}_j)\log \frac{1}{1+\mathrm{e}^{r_j^{(i)}}} \\ &= \sum_{j=1}^{d} \tilde{s}_j \log \frac{\mathrm{e}^{r_j^{(i)}}}{1+\mathrm{e}^{r_j^{(i)}}} - (1-\tilde{s}_j)\log(1+\mathrm{e}^{r_j^{(i)}}) \\ &= \sum_{j=1}^{d} \tilde{s}_j r_j^{(i)} - \log(1+\mathrm{e}^{r_j^{(i)}}) \\ &= \sum_{j=1}^{d} \tilde{s}_j r_j^{(i)} - \mathrm{softplus}(r_j^{(i)}). \end{aligned} \tag{8}$$

Therefore, we have

$$\log h_k(\tilde{s}) = \mathrm{logsumexp}\left(\sum_{j=1}^{d}\left(\tilde{s}_j r_j^{(i)} - \mathrm{softplus}(r_j^{(i)})\right) - \log k\right). \tag{9}$$

Putting all together, we arrive at the final expression

$$\mathcal{L}_k \approx \log p_\theta(x|s) - \sum_{i=1}^{d} \log 2 - \text{logsumexp}\left(\sum_{j=1}^{d}\left(\tilde{s}_j r_j^{(i)} - \text{softplus}(r_j^{(i)})\right) - \log k\right),$$

$$\text{where } s \sim q_\phi(s|x),\ \tilde{s} \sim h_k(\tilde{s}) = \frac{1}{k}\sum_{i=1}^{k} q_\phi(\tilde{s}|r^{(i)}),\ r^{(1\ldots k)} \sim q_\phi(r^{(1\ldots k)}|x). \tag{10}$$

Here, we end up the proof. □

References

[1] Lin Zheng, Qinliang Su, Dinghan Shen, and Changyou Chen. Generative semantic hashing enhanced via boltzmann machines. *arXiv preprint arXiv:2006.08858*, 2020.