

統計的機械学習レポート

工学部電子情報工学科 3 年 03190449 堀 紡希

6 月 18 日

1 レポート課題 1 データが複数ある場合

$q(z_{1:n}) = \prod_{i=1}^n q(z_i)$ と置いた時

$$\begin{aligned}\log p_{\theta}(x_{1:n}) &= \log \int p_{\theta}(x_{1:n}, z_{1:n}) dz_{1:n} \\ &= \log \int q(z_{1:n}) \frac{p_{\theta}(x_{1:n}, z_{1:n})}{q(z_{1:n})} dz_{1:n} \\ &\geq \int q(z_{1:n}) \log \frac{p_{\theta}(x_{1:n}, z_{1:n})}{q(z_{1:n})} dz_{1:n} \\ &= \int q(z_{1:n}) (\log p_{\theta}(x_{1:n}, z_{1:n}) - \log q(z_{1:n})) dz_{1:n} \\ &= \int \prod_{i=1}^n q(z_i) (\log \prod_{i=1}^n p_{\theta}(x_i, z_i) - \log q(z_{1:n})) dz_1 \dots dz_n \\ &= \int \prod_{i=1}^n q(z_i) (\sum_{i=1}^n \log p_{\theta}(x_i, z_i) - \log q(z_{1:n})) dz_1 \dots dz_n \\ &= \sum_{i=1}^n \int q(z_i) \log \frac{p_{\theta}(x_i|z_i)p(z_i)}{q(z_i)} dz_i \\ &= \sum_{i=1}^n L[q(z_i), \theta; x_i]\end{aligned}$$

より示された。

2 レポート課題 2 周辺尤度と KL 情報量の関係の導出

$$\begin{aligned}
& \log p(x_{1:n}|\eta) - L[q(z_{1:n})q(\theta); x_{1:n}] \\
&= \sum_z \int q(z_{1:n})q_\theta d\theta \log p(x_{1:n}|\eta) - \sum_z \int q(z_{1:n})q(\theta) \log \frac{p(x_{1:n}, z_{1:n}, \theta|\eta)}{q(z_{1:n})q(\theta)} d\theta \\
&= \sum_z \int q(z_{1:n})q_\theta \log \frac{p(x_{1:n}|\eta)q(z_{1:n})q(\theta)}{p(x_{1:n}, \theta|x_{1:n}, \eta)} d\theta \\
&= \sum_z \int q(z_{1:n})q(\theta) \log \frac{q(z)q(\theta)}{p(z, \theta|x, \eta)} d\theta \\
&= KL[q(z_{1:n})q(\theta)|p(z_{1:n}, \theta|x_{1:n}, \eta)]
\end{aligned}$$

より示された。