

Learning Structured Output Representation using Deep Conditional Generative Models

jojonki

January 2018

1 導出するよ

Sohn氏によるConditional VAEの論文, "Learning Structured Output Representation using Deep Conditional Generative Models", の補足資料でCVAEにおける変分下限を求めているが, こちらの導出をVAEの時と同様に丁寧にやってみる. 本論のCVAEでは入出力が同じでなく出力が y になっているため, VAEのときと違い入出力が異なる. またVAEでは式(1)のKLダイバージェンスを求めているのに対して, 本論のCVAEでは式(2)になる.

$$D_{KL}(q(z|x)||p(z|x)) \tag{1}$$

$$D_{KL}(q(z|x,y)||p(z|x,y)) \tag{2}$$

では、式(2)を展開していこう。基本的にはVAEのときの展開と同じであるが、3変数のベイズの定理を利用しているところだけが違う。元論文の補足資料にある式S1を目標に展開してみる。

$$\begin{aligned}
D_{KL}(q(z|x, y)||p(z|x, y)) &= \int q(z|x, y) \log \frac{q(z|x, y)}{p(z|x, y)} dz \\
&= \int q(z|x, y) \left\{ \log q(z|x, y) - \log p(z|x, y) \right\} dz \\
&= \int q(z|x, y) \log q(z|x, y) - \int q(z|x, y) \log p(z|x, y) dz \\
&\quad \text{第1項は期待値の定義そのもの.} \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \int q(z|x, y) \log p(z|x, y) dz \\
&\quad \text{第2項はベイズの定理で展開. } x, y \text{ は1つの塊としてみると普通のベイズの定理.} \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \int q(z|x, y) \log \frac{p(y, z|x)}{p(y, x)} dz \\
&\quad P(A, B) = P(A|B)P(B) \text{ の公式で、分母分子をそれぞれ展開.} \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \int q(z|x, y) \log \frac{p(y, z|x)p(x)}{p(y|x)p(x)} dz \\
&\quad p(x) \text{ が消える} \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \int q(z|x, y) \log \frac{p(y, z|x)}{p(y|x)} dz \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \int q(z|x, y) \log p(y, z|x) dz + \int (q(z|x, y) \log p(y|x)) dz \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \mathbb{E}_{q(z|x, y)} [\log p(y, z|x)] + \int q(z|x, y) \log p(y|x) dz \\
&\quad z \text{ で周辺化で積分消える} \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y)] - \mathbb{E}_{q(z|x, y)} [\log p(y, z|x)] + \log p(y|x) \\
&= \mathbb{E}_{q(z|x, y)} [\log q(z|x, y) - \log p(y, z|x)] + \log p(y|x)
\end{aligned} \tag{3}$$

$\log p(y|x)$ を左辺に持ってきて整理すると、論文の式S1になる。

$$\log p(y|x) = D_{KL}(q(z|x, y)||p(z|x, y)) + \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log p(y, z|x)]$$

KLダイバージェンスは0以上の値なので変分下限を適用してS2に。

$$\geq \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log p(y, z|x)]$$

先ほどと同様に2項目にベイズの定理を適用して変形していく。

$$= \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log \frac{p(y, x, z)}{p(x)}]$$

$$= \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log \frac{p(y|x, z)p(x, z)}{p(x)}]$$

$$= \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log \frac{p(y|x, z)p(z|x)p(x)}{p(x)}]$$

S3相当。

$$= \mathbb{E}_{q(z|x, y)}[-\log q(z|x, y) + \log p(y|x, z)p(z|x)]$$

$$= \mathbb{E}_{q(z|x, y)}[-\log \frac{q(z|x, y)}{p(z|x)} + \log p(y|x, z)]$$

$$= -\mathbb{E}_{q(z|x, y)}[\log \frac{q(z|x, y)}{p(z|x)}] + \mathbb{E}_{q(z|x, y)}[\log p(y|x, z)]$$

$$= -\int q(z|x, y) \log \frac{q(z|x, y)}{p(z|x)} dz + \mathbb{E}_{q(z|x, y)}[\log p(y|x, z)]$$

第1項はKLダイバージェンスの定義そのもの。これにより式S4を得る。これがCVAEのELBO。

$$= -D_{KL}(q(z|x, y)||p(z|x)) + \mathbb{E}_{q(z|x, y)}[\log p(y|x, z)]$$

(4)

ELBOの再構築ロスはVAEのときと同様にL個のサンプリングの平均で代替し、論文中の式(5)を得ることを確認できる。

$$\tilde{\mathcal{L}}_{\text{CVAE}(x, y; \theta, \phi)} = -D_{KL}(q(z|x, y)||p(z|x)) + \frac{1}{L} \sum_{l=1}^L \log p(y|x, z^{(l)}) \quad (5)$$

参考までにVAEのときの目的関数も書いておく。

$$\tilde{\mathcal{L}}_{\text{VAE}(x; \theta, \phi)} = -D_{KL}(q(z|x)||p(z)) + \frac{1}{L} \sum_{l=1}^L \log p(x|z^{(l)}) \quad (6)$$