# Variational Auto Encoder

jojonki

January 2018

## 1　導出するよ

Kingma氏のVAEの論文で(https://arxiv.org/pdf/1312.6114.pdf)，Appendix Bに多変量正規分布間におけるKLダイバージェンスの計算の式が載っていたが，途中式が省かれていたので導出を試みた．

変分下限の正則化項の導出を目標にこの資料を作ったが，せっかくなので大本の変分下限の下りも少しだけ触れる． $p(z|x)$の近似分布として，$q(z|x)$を考え，$D_{KL}(q(z|x)||p(z|x))$を考えるわけだが，変分下限を利用すると

$$D_{KL}(q(z|x)||p(z|x)) = \int q(z|x) \log \frac{q(z|x)}{p(z|x)}$$

ここの導出は目標でないのでnzwさんの導出を見てね https://nzw0301.github.io/notes/vae.pdf

$$= \log p(x) + D_{KL}(q(z|x)||p(z)) - \int q(z|x) \log p(x|z) dz \tag{1}$$

これを整理すると下記のようになる，

$$\log p(x) - D_{KL}(q(z|x)||p(z)) = \int q(z|x) \log p(z|x) dz - D_{KL}(q(z|x)||p(z)) \tag{2}$$

KLダイバージェンスは0以上であるわけだから，左辺のKLダイバージェンスを取り除くと下記の不等号が成り立つ．

$$\log p(x) \geq \int q(z|x) \log p(z|x) - D_{KL}(q(z|x)||p(z)) \tag{3}$$

このときの右辺がEvidence Lower BOund (ELBO)と呼ばれ，これを下限を最大化することで$p(x)$を高めることができる．第1項が復元誤差(Reconstruction error)，第2項が正則化項となる．

$$\mathcal{L} = \int q(z|x) \log p(z|x) - D_{KL}(q(z|x)||p(x|z)) \tag{4}$$

### 1.1　復元誤差の導出

略

## 1.2 正則化項の導出

$$
\begin{aligned}
D_{KL}(q(z|x)||p(z)) &= \int q(z|x)\log\frac{q(z|x)}{p(z)}dz \\
&= \int q(z|x)\{\log q(z|x) - \log p(z)\}dz \\
&= \int q(z|x)\log q(z|x)dz - \int q(z|x)\log p(z)dz
\end{aligned} \tag{5}
$$

第 1 項と第 2 項をそれぞれ求めたいと思う． $p(z)$はVAEの設定から $\sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ .であり， $q(z|x)$は$\sim$ $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$となる．

### 第1項

$$
\begin{aligned}
\int q(z|x)\log q(z|x)dz &= \int \mathcal{N}(z;\mu,\sigma^2)\log\mathcal{N}(z;\mu,\sigma^2)dz \\
&= \sum_j^J E_{q(z_j)}\Big[\log\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp\big(-\frac{(z_j-\mu_j)^2}{2\sigma_j^2})\Big] \\
&= \sum_j^J E_{q(z_j)}\Big[-\frac{1}{2}\log(2\pi\sigma_j^2) - \frac{(z_j-\mu_j)^2}{2\sigma_j^2}\Big] \\
&= \sum_j^J\Big\{-\frac{1}{2}\log(2\pi\sigma_j^2) - E_{q(z_j)}\Big[\frac{(z_j-\mu_j)^2}{2\sigma_j^2}\Big]\Big\} \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^J\log\sigma_j^2 - \frac{1}{2}\sum_{j=1}^J E_{q(z_j)}\Big[\frac{(z_j-\mu_j)^2}{\sigma_j^2}\Big] \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^J\log\sigma_j^2 - \frac{1}{2}\sum_{j=1}^J\frac{1}{\sigma_j^2}E_{q(z_j)}\Big[(z_j-\mu_j)^2\Big] \\
&\quad \text{第3項の右側は分散の定義そのもの.} \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^J\log\sigma_j^2 - \frac{1}{2}\sum_{j=1}^J\frac{1}{\sigma_j^2}\sigma_j^2 \\
&= -\frac{J}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^J(\log\sigma_j^2 + 1)
\end{aligned} \tag{6}
$$

第2項

$$
\begin{aligned}
\int q(z|x)\log p(z)dz &= \int \mathcal{N}(z;\mu,\sigma^2)\log \mathcal{N}(z;\mathbf{0},\boldsymbol{I})dz \\
&= \sum_{j}^{J} E_{q(z_j)}\Big[\log \frac{1}{\sqrt{2\pi}}\exp\big(-\frac{z_j^2}{2}\big)\Big] \\
&= \sum_{j}^{J} E_{q(z_j)}\Big[-\frac{1}{2}\log(2\pi)-\frac{z_j^2}{2}\Big] \\
&= \sum_{j}^{J}\Big\{-\frac{1}{2}\log(2\pi)-E_{q(z_j)}\Big[\frac{z_j^2}{2}\Big]\Big\} \\
&= -\frac{J}{2}\log(2\pi)-\frac{1}{2}\sum_{j=1}^{J}E_{q(z_j)}\Big[z_j^2\Big] \\
&\quad \text{ここで第 2 項に分散の公式を使う．} \quad \sigma^2 = E[X^2]-\mu^2 \\
&= -\frac{J}{2}\log(2\pi)-\frac{1}{2}\sum_{j=1}^{J}(\sigma_j^2+\mu_j^2)
\end{aligned} \tag{7}
$$

式(5)の第 1 項第 2 項が求まったので，式(5)をまとめると下記のようになり，論文の結果と一致することになる．(論文ではKLダイバージェンスに符号がついているので注意)

$$
\begin{aligned}
D_{KL}(q(z|x)||p(z)) &= \int q(z|x)\log\frac{q(z|x)}{p(z)}dz \\
&= \int q(z|x)\{\log q(z|x)-\log p(z)\}dz \\
&= \int q(z|x)\log q(z|x)dz - \int q(z|x)\log p(z)dz \\
&= \left(-\frac{J}{2}\log(2\pi)-\frac{1}{2}\sum_{j=1}^{J}(\log\sigma_j^2+1)\right)-\left(-\frac{J}{2}\log(2\pi)-\frac{1}{2}\sum_{j=1}^{J}(\sigma_j^2+\mu_j^2)\right) \\
&= -\frac{1}{2}\sum_{j=1}^{J}\left(1+\log\sigma^2-\sigma_j^2-\mu_j^2\right)
\end{aligned} \tag{8}
$$