

(6.10) について <https://data-analytics.fun/2021/04/14/understanding-vae/>の”Variational Auto-Encoder(VAE) の概要”の最後の式との対応を考える。

<https://data-analytics.fun/2021/04/14/understanding-vae/>の”Variational Auto-Encoder(VAE) の概要”の最後の式は、本の書き方に近づけると、

$$\mathcal{L}(x, \psi) = \mathbb{E}_{q(Z; X, \psi)}[\ln p(X|Z)] - D_{KL}(q(Z; X, \psi) || p(Z)) \quad (1)$$

ページで θ は本で W に相当する。学習するパラメータにはなるが、これは、ページのコードでは確率変数ではない。(P.148 の注記*1 参照)

一方、(6.10) は W を確率変数としない場合、

$$\mathcal{L}(x, \psi) = \mathbb{E}_{q(Z; X, \psi)}[\ln p(X, Z)] - \mathbb{E}_{q(Z; X, \psi)}[\ln q(Z; X, \psi)] \quad (2)$$

となる。これを変形する。

$$\begin{aligned} \mathcal{L}(x, \psi) &= \mathbb{E}_{q(Z; X, \psi)}[\ln p(X, Z)] - \mathbb{E}_{q(Z; X, \psi)}[\ln q(Z; X, \psi)] = \mathbb{E}_{q(Z; X, \psi)}[\ln p(X|Z)p(Z)] - \mathbb{E}_{q(Z; X, \psi)}[\ln q(Z; X, \psi)] \\ &= \mathbb{E}_{q(Z; X, \psi)}[\ln p(X|Z)] + \mathbb{E}_{q(Z; X, \psi)}[\ln p(Z)] - \mathbb{E}_{q(Z; X, \psi)}[\ln q(Z; X, \psi)] = \mathbb{E}_{q(Z; X, \psi)}[\ln p(X|Z)] - D_{KL}(q(Z; X, \psi) || p(Z)) \end{aligned} \quad (3)$$

なお、

$$D_{KL}(q(Z; X, \psi) || p(Z)) = \mathbb{E}_{q(Z; X, \psi)}[\ln q(Z; X, \psi)] - \mathbb{E}_{q(Z; X, \psi)}[\ln p(Z)] = \int q(Z; X, \psi) \ln \frac{q(Z; X, \psi)}{p(Z)} dZ \quad (4)$$