

式 (5.16) を見直してみる。

式 (4.21) の上を参考にすると、

$$-\mathcal{U}(z) = \ln \tilde{p}(z) \quad (\ln p(z) \propto \ln \tilde{p}(z)) \quad (1)$$

このとき、 $p(z)$ から z をサンプリングしたい。

さて、今回、P.117 にあるように $p(W|Y, X)$ から、 W をサンプリングしたい。

よって、1 倍は比例でもあるので、 $\tilde{p}(W|Y, X) = p(W|Y, X)$ として、

$$\mathcal{U}(W) = -\ln \tilde{p}(W|Y, X) = -\ln p(W|Y, X) = -\ln \frac{p(W, Y|X)}{p(Y|X)} = -\ln \frac{p(Y|W, X)p(W|X)}{p(Y|X)} = -\ln \frac{p(Y|W, X)p(W)}{p(Y|X)} \quad (2)$$

ここで、(5.1) を考えると、 W と X は独立で、 $p(W|X) = p(W)$ となることに注意する。

(4.16),(4.17) を考慮すると、 $\mathcal{U}(W)$ は微分だけで考慮されている。すると、(2) で $p(Y|X)$ は W を変数とすると、定数となり、無視できる。

よって、(2) は、以下のように考えられ、(5.16) が求まる。

$$\mathcal{U}(W) = -\ln p(Y|W, X)p(W) = -\{\ln p(Y|W, X) + \ln p(W)\} \quad (3)$$

なお、(4.16) をなどを考慮すると、 $\frac{\partial \mathcal{U}}{\partial w_i}$ がわかれば良い。((4.14),(4.15) は一部、偏微分になるはず。)

(2) の最初の等号の偏微分を考慮すると、(5.8) が利用できることがわかる。

$$\frac{\partial \mathcal{U}}{\partial w_i} = -\frac{\partial \mathcal{U}}{\partial w_i} = \frac{\partial \ln p(W|Y, X)}{\partial w_i} = \frac{1}{\sigma_y^2} \frac{\partial}{\partial w_i} E(W) + \frac{1}{\sigma_w^2} \frac{\partial}{\partial w_i} \Omega_{L2}(W) \quad (4)$$

$$\nabla_w \mathcal{U}(W) = \frac{1}{\sigma_y^2} \nabla_w E(W) + \frac{1}{\sigma_w^2} W \quad (5)$$