

式 (6.45) の展開を考えてみる。

$$\begin{aligned}
& \mathcal{L}(\psi, \xi) = \ln p(X) - D_{KL}[q(Z, \theta; X, \psi, \xi) | p(Z, \theta | X)] \text{ (式 (6.10))} \\
& = \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln p(X)] + \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln p(Z, \theta | X)] - \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln q(Z, \theta; X, \psi, \xi)] \\
& \quad (p(X) \text{ は定数で期待値は } p(X) \text{ のまま。} D_{KL} \text{ を展開}) \\
& = \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln(p(X)p(Z, \theta | X))] - \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln q(Z, \theta; X, \psi, \xi)] \text{ (期待値をまとめる。)} \\
& \quad = \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln p(X, Z, \theta)] - \mathbb{E}_{q(Z, \theta; X, \psi, \xi)}[\ln q(Z, \theta; X, \psi, \xi)] \text{ (ベイズの定理)} \\
& \quad = \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln p(X, Z, \theta)] - \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln q(Z, \theta; X, \psi, \xi)] \text{ (式 (6.44))} \\
& = \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln(p(\theta) \prod_n (p(x_n | z_n, \theta) p(z_n, \theta)))] - \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln(q_\xi(\theta) \prod_n q_\psi(z_n | x_n, \theta))] \text{ (式 (6.39), (6.44))} \\
& \quad = \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln(p(\theta) \prod_n p(x_n, z_n | \theta))] - \mathbb{E}_{\prod_n q_\psi(z_n | x_n, \theta) q_\xi(\theta)}[\ln(q_\xi(\theta) \prod_n q_\psi(z_n | x_n, \theta))] \text{ (ベイズの定理)} \\
& = \mathbb{E}_{q_\xi(\theta)}[\ln p(\theta)] + \mathbb{E}_{q_\xi(\theta) \prod_n q_\psi(z_n | x_n, \theta)}[\ln(\prod_n p(x_n, z_n | \theta))] - \mathbb{E}_{q_\xi(\theta)}[\ln q_\xi(\theta)] - \mathbb{E}_{q_\xi(\theta) \prod_n q_\psi(z_n | x_n, \theta)}[\ln(\prod_n q_\psi(z_n | x_n, \theta))] \\
& \quad (\text{変数が } \theta \text{ のみの部分は } q_\psi \text{ によらない。} z_n \text{ で期待値を取ると 1 になるので、} q_\psi \text{ は消える。}) \\
& = \mathbb{E}_{q_\xi(\theta)}[\ln p(\theta) - \ln q_\xi(\theta)] + \mathbb{E}_{q_\xi(\theta) \prod_n q_\psi(z_n | x_n, \theta)}[\ln \prod_n p(x_n, z_n | \theta) - \ln \prod_n q_\psi(z_n | x_n, \theta)] \\
& \quad (\text{同じ確率分布の期待値をまとめる。}) \\
& = \mathbb{E}_{q_\xi(\theta)}[\ln p(\theta) - \ln q_\xi(\theta)] + \sum_n \mathbb{E}_{q_\xi(\theta) \prod_m q_\psi(z_m | x_m, \theta)}[\ln p(x_n, z_n | \theta) - \ln q_\psi(z_n | x_n, \theta)] \\
& \quad (\ln \text{ の中の } \prod \text{ をばらす。確率分布の引数が } m \text{ になっていることに注意。}) \\
& = \mathbb{E}_{q_\xi(\theta)}[\ln p(\theta) - \ln q_\xi(\theta)] + \sum_n \mathbb{E}_{q_\xi(\theta) q_\psi(z_n | x_n, \theta)}[\ln p(x_n, z_n | \theta) - \ln q_\psi(z_n | x_n, \theta)] \\
& \quad (n \neq m \text{ のときは期待値を変化させないので、消去する。}) \\
& \quad (1)
\end{aligned}$$

となり、(6.45) が求まる。