Denoising Diffusion Probabilistic Models の変分推論

正田 備也

masada@rikkyo.ac.jp

Contents

周辺尤度 (marginal likelihood)

以下の結合分布 (joint distribution) を持つ確率モデルを考える。

$$p_\theta(\mathbf{x}_{0:T}) = p_\theta(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) \tag{1}$$

ただし、$\mathbf{x}_0$ は観測データ、$\{\mathbf{x}_t : t = 1, \ldots, T\}$ は潜在的な確率変数である。式 (1) が表すように、$\mathbf{x}_{t-1}$ の分布は $\mathbf{x}_t$ だけに条件付けられている。そして、$\mathbf{x}_0$ の対数周辺尤度は次のように書ける。

$$\log p_\theta(\mathbf{x}_0) = \log \int p_\theta(\mathbf{x}_{0:T}) d\mathbf{x}_{1:T} \tag{2}$$

Contents

ELBO

Jensen の不等式は対数周辺尤度の下界を次のように与える。

$$\begin{aligned}\log p_\theta(\mathbf{x}_0) &= \log \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\frac{p_\theta(\mathbf{x}_{0:T})}{q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)}d\mathbf{x}_{1:T} \\ &\geq \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_{0:T})}{q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)}d\mathbf{x}_{1:T} \\ &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_{0:T})}{q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)}d\mathbf{x}_{1:T} \equiv L_{\text{VLB}} \qquad (3)\end{aligned}$$

ただし $q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)$ は変分事後分布で、VAE 同様、観測データ $\mathbf{x}_0$ に条件付けられている。そしてモデルは、観測データの集合 $\mathcal{X} \equiv \{\mathbf{x}_0^{(1)}, \ldots, \mathbf{x}_0^{(N)}\}$ の上で、amortized な仕方で訓練される。

マルコフ性の仮定

条件付き分布の定義より、

$$
\begin{aligned}
q_\psi(\mathbf{x}_2|\mathbf{x}_1,\mathbf{x}_0)q_\psi(\mathbf{x}_1|\mathbf{x}_0) &= \frac{q_\psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_1,\mathbf{x}_0)}\frac{q_\psi(\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_0)} = \frac{q_\psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_0)} \\
&= q_\psi(\mathbf{x}_2,\mathbf{x}_1|\mathbf{x}_0) = q_\psi(\mathbf{x}_{1:2}|\mathbf{x}_0) \\
q_\psi(\mathbf{x}_3|\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)q_\psi(\mathbf{x}_2|\mathbf{x}_1,\mathbf{x}_0)q_\psi(\mathbf{x}_1|\mathbf{x}_0) &= \frac{q_\psi(\mathbf{x}_3,\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)}\frac{q_\psi(\mathbf{x}_2,\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_1,\mathbf{x}_0)}\frac{q_\psi(\mathbf{x}_1,\mathbf{x}_0)}{q_\psi(\mathbf{x}_0)} \\
&= q_\psi(\mathbf{x}_3,\mathbf{x}_2,\mathbf{x}_1|\mathbf{x}_0) = q_\psi(\mathbf{x}_{1:3}|\mathbf{x}_0) \\
&\cdots \\
q_\psi(\mathbf{x}_1|\mathbf{x}_0)\prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\ldots,\mathbf{x}_1,\mathbf{x}_0) &= q_\psi(\mathbf{x}_T,\ldots,\mathbf{x}_1|\mathbf{x}_0) = q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \qquad (4)
\end{aligned}
$$

ここで、$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\ldots,\mathbf{x}_1,\mathbf{x}_0) = q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0)$ が $t = 2,\ldots,T$ について成り立つと仮定することによって、変分事後分布を単純化する。

このマルコフ性の仮定により、$q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)$ は、次のように分解できることになる。

$$q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) = q_\psi(\mathbf{x}_1|\mathbf{x}_0) \prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0) \tag{5}$$

このとき、式 (3) の変分下界 L_{VLB} は、以下のように書き直せる。

$$\begin{aligned} L_{\mathsf{VLB}} &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{0:T})}{q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)} d\mathbf{x}_{1:T} \\ &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_1|\mathbf{x}_0) \prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0)} d\mathbf{x}_{1:T} \\ &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log p_\theta(\mathbf{x}_T) d\mathbf{x}_{1:T} + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \sum_{t=2}^{T} \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0)} d\mathbf{x}_{1:T} \\ &\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_0|\mathbf{x}_1)}{q_\psi(\mathbf{x}_1|\mathbf{x}_0)} d\mathbf{x}_{1:T} \end{aligned} \tag{6}$$

式 (6) に現れる $q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0)$ についてベイズ則を使うと、次を得る。

$$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0) = \frac{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)} \tag{7}$$

(なぜこんなことをするのかは、p. 19 で明らかになる。)
この式 (7) にもとづいて、式 (6) に現れる $q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0)$ を $\frac{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)}$ で置き換えると、変分下界 L_{VLB} は以下のように書き換えられる。

$$\begin{aligned} L_{\text{VLB}} &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log p_\theta(\mathbf{x}_T)d\mathbf{x}_{1:T} + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\sum_{t=2}^{T}\log\frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)}d\mathbf{x}_{1:T} \\ &\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\sum_{t=2}^{T}\log\frac{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q_\psi(\mathbf{x}_t|\mathbf{x}_0)}d\mathbf{x}_{1:T} + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_0|\mathbf{x}_1)}{q_\psi(\mathbf{x}_1|\mathbf{x}_0)}d\mathbf{x}_{1:T} \end{aligned} \tag{8}$$

(次のページに続く。)

$$
\begin{aligned}
L_{\mathsf{VLB}} &= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log p_\theta(\mathbf{x}_T)d\mathbf{x}_{1:T} + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\sum_{t=2}^{T}\log\frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)}d\mathbf{x}_{1:T} \\
&\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{q_\psi(\mathbf{x}_1|\mathbf{x}_0)\cancel{q_\psi(\mathbf{x}_2|\mathbf{x}_0)}\cdots\cancel{q_\psi(\mathbf{x}_{T-1}|\mathbf{x}_0)}}{\cancel{q_\psi(\mathbf{x}_2|\mathbf{x}_0)}\cancel{q_\psi(\mathbf{x}_2|\mathbf{x}_0)}\cdots q_\psi(\mathbf{x}_T|\mathbf{x}_0)}d\mathbf{x}_{1:T} \\
&\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_0|\mathbf{x}_1)}{q_\psi(\mathbf{x}_1|\mathbf{x}_0)}d\mathbf{x}_{1:T} \\
&= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log p_\theta(\mathbf{x}_T)d\mathbf{x}_{1:T} + \sum_{t=2}^{T}\int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)}d\mathbf{x}_{1:T} \\
&\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{\cancel{q_\psi(\mathbf{x}_1|\mathbf{x}_0)}}{q_\psi(\mathbf{x}_T|\mathbf{x}_0)}d\mathbf{x}_{1:T} + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_0|\mathbf{x}_1)}{\cancel{q_\psi(\mathbf{x}_1|\mathbf{x}_0)}}d\mathbf{x}_{1:T} \\
&= \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_T)}{q_\psi(\mathbf{x}_T|\mathbf{x}_0)}d\mathbf{x}_{1:T} + \sum_{t=2}^{T}\int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log\frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)}d\mathbf{x}_{1:T} \\
&\quad + \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)\log p_\theta(\mathbf{x}_0|\mathbf{x}_1)d\mathbf{x}_{1:T} \equiv L_T + \sum_{t=2}^{T}L_{t-1} + L_0 \qquad (9)
\end{aligned}
$$

ここで、式 (7) より

$$
\begin{aligned}
q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0)q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_{t-2},\mathbf{x}_0) &= \frac{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)q_\psi(\mathbf{x}_t|\mathbf{x}_0)q_\psi(\mathbf{x}_{t-2}|\mathbf{x}_{t-1},\mathbf{x}_0)q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)q_\psi(\mathbf{x}_{t-2}|\mathbf{x}_0)} \\
&= \frac{q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-2}|\mathbf{x}_0)}q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)q_\psi(\mathbf{x}_{t-2}|\mathbf{x}_{t-1},\mathbf{x}_0) \qquad (10)
\end{aligned}
$$

同様に考えて、

$$
\prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0) = \frac{q_\psi(\mathbf{x}_T|\mathbf{x}_0)}{q_\psi(\mathbf{x}_1|\mathbf{x}_0)}\prod_{t=2}^{T} q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0) \qquad (11)
$$

両辺に $q_\psi(\mathbf{x}_1|\mathbf{x}_0)$ を掛けて

$$
q_\psi(\mathbf{x}_1|\mathbf{x}_0)\prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{x}_0) = q_\psi(\mathbf{x}_T|\mathbf{x}_0)\prod_{t=2}^{T} q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0) \qquad (12)
$$

式 (5) より、これは $q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0)$ に等しい。

式 (12) より、式 (9) の L_{t-1} は下のように書き換えられる。

$$
\begin{aligned}
L_{t-1} &\equiv \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:T} \\
&= \int \left(q_\psi(\mathbf{x}_T|\mathbf{x}_0) \prod_{t' \neq t} q_\psi(\mathbf{x}_{t'-1}|\mathbf{x}_{t'}, \mathbf{x}_0) \right) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:T} \qquad (13)
\end{aligned}
$$

ここで

$$
\begin{aligned}
q_\psi(\mathbf{x}_T|\mathbf{x}_0) q_\psi(\mathbf{x}_{T-1}|\mathbf{x}_T, \mathbf{x}_0) &= \frac{q_\psi(\mathbf{x}_T, \mathbf{x}_0) q_\psi(\mathbf{x}_{T-1}, \mathbf{x}_T, \mathbf{x}_0)}{q_\psi(\mathbf{x}_0) q_\psi(\mathbf{x}_T, \mathbf{x}_0)} = \frac{q_\psi(\mathbf{x}_{T-1}, \mathbf{x}_T, \mathbf{x}_0)}{q_\psi(\mathbf{x}_0)} \\
&= q_\psi(\mathbf{x}_{T-1}, \mathbf{x}_T|\mathbf{x}_0) \qquad (14)
\end{aligned}
$$

であるから、$\mathbf{x}_T$ を積分消去すると、以下を得る。

$$
\begin{aligned}
L_{t-1} &= \int \left(q_\psi(\mathbf{x}_{T-1}, \mathbf{x}_T|\mathbf{x}_0) \prod_{t' \neq t \,\wedge\, t' < T} q_\psi(\mathbf{x}_{t'-1}|\mathbf{x}_{t'}, \mathbf{x}_0) \right) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:T} \\
&= \int \left(q_\psi(\mathbf{x}_{T-1}|\mathbf{x}_0) \prod_{t' \neq t \,\wedge\, t' < T} q_\psi(\mathbf{x}_{t'-1}|\mathbf{x}_{t'}, \mathbf{x}_0) \right) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:T-1}
\end{aligned}
$$

同様に、$\mathbf{x}_{T-1}, \ldots, \mathbf{x}_{t+1}$ を積分消去すると、以下を得る。

$$L_{t-1} = \int \Big(q_\psi(\mathbf{x}_t|\mathbf{x}_0) \prod_{t'=2}^{t-1} q_\psi(\mathbf{x}_{t'-1}|\mathbf{x}_{t'}, \mathbf{x}_0) \Big) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:t} \tag{15}$$

ここで、再び式 (12) を使うと

$$\begin{aligned}
L_{t-1} &= \int \Big(\frac{q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)} q_\psi(\mathbf{x}_1|\mathbf{x}_0) \prod_{t'=2}^{t-1} q_\psi(\mathbf{x}_{t'}|\mathbf{x}_{t'-1}, \mathbf{x}_0) \Big) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{1:t} \\
&= \int \Big(\frac{q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)} q_\psi(\mathbf{x}_2|\mathbf{x}_0) \prod_{t'=3}^{t-1} q_\psi(\mathbf{x}_{t'}|\mathbf{x}_{t'-1}, \mathbf{x}_0) \Big) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{2:t} \\
&\cdots \\
&= \int \Big(\frac{q_\psi(\mathbf{x}_t|\mathbf{x}_0)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0)} q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_0) \Big) q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{t-1:t} \\
&= \int q_\psi(\mathbf{x}_t|\mathbf{x}_0) \Big(\int q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} d\mathbf{x}_{t-1} \Big) d\mathbf{x}_t \\
&\equiv -\mathbb{E}_{q_\psi(\mathbf{x}_t|\mathbf{x}_0)} \big[D_{\mathsf{KL}}(q_\psi(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \parallel p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)) \big]
\end{aligned} \tag{16}$$

Contents

変分事後分布の一つの設定方法

変分事後分布 $q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) = q_\psi(\mathbf{x}_1|\mathbf{x}_0)\prod_{t=2}^{T} q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0)$ は、$\psi \equiv \{\alpha_1, \ldots, \alpha_T\}$ をパラメータとする以下のような多変量正規分布の積として構成されていると仮定する。

$$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I}) \quad \text{for } t = 2, \ldots, T \tag{17}$$

こう仮定すると、$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0) = q_\psi(\mathbf{x}_t|\mathbf{x}_{t-1})$ となり、$\mathbf{x}_1$ 以外は $\mathbf{x}_0$ に依存しなくなる。

Appendix の式 (34) より、$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-2})$ は、下のように書き換えられる。

$$q_\psi(\mathbf{x}_t|\mathbf{x}_{t-2}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t\alpha_{t-1}}\mathbf{x}_{t-2}, (1-\alpha_t\alpha_{t-1})\mathbf{I}) \tag{18}$$

同じ議論を繰り返すと、$q_\psi(\mathbf{x}_t|\mathbf{x}_0)$ は、下のように書き換えられる。

$$q_\psi(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}) \tag{19}$$

ただし、$\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$ である。この $q_\psi(\mathbf{x}_t|\mathbf{x}_0)$ からは、簡単にサンプルを得られる（式 (24) を参照）。よって、式 (16) の期待値は、モンテカルロ近似できる。

ψ を自由パラメータとみなすことにし、これ以降、ψ を notations から脱落させることにする。

式 (17) と式 (19) より、式 (16) に現れる $q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$ は、以下のように書き換えられる。

$$
\begin{aligned}
q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) &= \frac{q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0)q(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} \quad \text{(based on Eq. (7))} \\
&\propto \exp\Bigg(-\frac{1}{2}\Big(\frac{(\mathbf{x}_t - \sqrt{\alpha_t}\mathbf{x}_{t-1})^2}{1-\alpha_t} + \frac{(\mathbf{x}_{t-1} - \sqrt{\bar{\alpha}_{t-1}}\mathbf{x}_0)^2}{1-\bar{\alpha}_{t-1}} - \frac{(\mathbf{x}_t - \sqrt{\bar{\alpha}_{t1}}\mathbf{x}_0)^2}{1-\bar{\alpha}_t}\Big)\Bigg) \\
&\propto \exp\Bigg(-\frac{1}{2}\Big((\frac{\alpha_t}{1-\alpha_t} + \frac{1}{1-\bar{\alpha}_{t-1}})\mathbf{x}_{t-1}^2 - 2(\frac{\sqrt{\alpha_t}}{1-\alpha_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}\mathbf{x}_0)\mathbf{x}_{t-1}\Big)\Bigg)
\end{aligned}
\tag{20}
$$

つまり、$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$ は正規分布であることが分かる。
そこで、その平均を $\tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0)$、分散を $\tilde{\beta}_t$ と書くことにする。つまり

$$
q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \equiv \mathcal{N}(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\beta}_t) \tag{21}
$$

と設定する。

すると、以下を得る。

$$\tilde{\beta}_t = 1/\left(\frac{\alpha_t}{1-\alpha_t} + \frac{1}{1-\bar{\alpha}_{t-1}}\right) = \frac{(1-\alpha_t)(1-\bar{\alpha}_{t-1})}{\alpha_t - \alpha_t\bar{\alpha}_{t-1} + 1 - \alpha_t} = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}(1-\alpha_t) \tag{22}$$

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0) &= \left(\frac{\sqrt{\alpha_t}}{1-\alpha_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}\mathbf{x}_0\right) \Big/ \left(\frac{\alpha_t}{1-\alpha_t} + \frac{1}{1-\bar{\alpha}_{t-1}}\right) \\
&= \left(\frac{\sqrt{\alpha_t}}{1-\alpha_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}\mathbf{x}_0\right)\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}(1-\alpha_t) \\
&= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\mathbf{x}_0
\end{aligned} \tag{23}$$

式 (19) にもとづくと、$\mathbf{x}_t$ は、次のように reparameterize できる。

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon} \quad \text{for } \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{24}$$

よって

$$\mathbf{x}_0 = \frac{\mathbf{x}_t}{\sqrt{\bar{\alpha}_t}} - \frac{\sqrt{1-\bar{\alpha}_t}}{\sqrt{\bar{\alpha}_t}}\boldsymbol{\epsilon} \quad \text{for } \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{25}$$

式 (25) を式 (23) に代入することで、$\tilde{\boldsymbol{\mu}}$ を以下のように書き換えることができる。

$$
\begin{aligned}
\tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \boldsymbol{\epsilon}) &= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\left(\frac{\mathbf{x}_t}{\sqrt{\bar{\alpha}_t}} - \frac{\sqrt{1-\bar{\alpha}_t}}{\sqrt{\bar{\alpha}_t}}\boldsymbol{\epsilon}\right) \\
&= \left(\frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} + \frac{1-\alpha_t}{(1-\bar{\alpha}_t)\sqrt{\alpha_t}}\right)\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}\sqrt{\alpha_t}}\boldsymbol{\epsilon} \\
&= \frac{1}{\sqrt{\alpha_t}}\left(\left(\frac{\alpha_t(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} + \frac{1-\alpha_t}{1-\bar{\alpha}_t}\right)\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}\right) \\
&= \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}\right) \qquad (26)
\end{aligned}
$$

Contents

観測データのモデリング

ここで初めて、生成モデルの詳細を以下のように指定する。

$$p_\theta(\mathbf{x}_T) = \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I}) \tag{27}$$

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)) \tag{28}$$

ここで、$\boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t) = \sigma_t^2\mathbf{I}$ と仮定する（**[?]** を参照）。

式 (21) と式 (28) より、式 (16) にある KL 情報量は、一つの正規分布から別の正規分布への KL 情報量であると分かる。したがって、式 (16) の L_{t-1} は、以下のように書き直せる。[1]

$$L_{t-1} = -\mathbb{E}_{q(\mathbf{x}_t|\mathbf{x}_0)}\left[\frac{1}{2\sigma_t^2}\left\|\tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \boldsymbol{\epsilon}) - \boldsymbol{\mu}_\theta(\mathbf{x}_t, t)\right\|^2\right] + const. \tag{29}$$

[1] https://scoste.fr/posts/dkl_gaussian/

式 (26) を使うと、L_{t-1} は、さらに、以下のように書き直せる。

$$L_{t-1} = -\mathbb{E}_{q(\mathbf{x}_t|\mathbf{x}_0)}\left[\frac{1}{2\sigma_t^2}\left\|\frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}\right) - \boldsymbol{\mu}_\theta(\mathbf{x}_t, t)\right\|^2\right] + const. \tag{30}$$

ここで、$\boldsymbol{\mu}_\theta(\mathbf{x}_t, t)$ を、次のように parameterize することを考える（[?] を参照)。

$$\boldsymbol{\mu}_\theta(\mathbf{x}_t, t) = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) \tag{31}$$

ただし、$\boldsymbol{\epsilon}_\theta$ は関数である。式 (31) の parameterization を使うことで L_{t-1} が以下のように書き換えられることより、この関数 $\boldsymbol{\epsilon}_\theta$ は、$\boldsymbol{\epsilon}$ を予測する役割を果たすと言える。

$$L_{t-1} = -\mathbb{E}_{\boldsymbol{\epsilon}\sim\mathcal{N}(\mathbf{0},\mathbf{I})}\left[\frac{(1-\alpha_t)^2}{2\sigma_t^2(1-\bar{\alpha}_t)}\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t)\|^2\right] + const. \tag{32}$$

なお、上の式では、$\mathbf{x}_t$ を式 (24) にもとづいて $\mathbf{x}_0$ の式で置き換えている。これにともなって、外側の期待値も、標準正規分布に関する期待値に書き換えてある。(つまり、[?] の Algorithm 1 の 4 行目は、上式の期待値をモンテカルロ近似するためのサンプリングになっている。)

さて、次に、式 (9) の L_T を考える。

$$L_T \equiv \int q_\psi(\mathbf{x}_{1:T}|\mathbf{x}_0) \log \frac{p_\theta(\mathbf{x}_T)}{q_\psi(\mathbf{x}_T|\mathbf{x}_0)} d\mathbf{x}_{1:T} \tag{33}$$

ノイズ分布 $p_\theta(\mathbf{x}_T)$ と近似事後分布 $q_\psi(\mathbf{x}_T|\mathbf{x}_0)$ は、trainable なパラメータを持たない。したがって、L_T は定数と見なせる。

最後に、式 (9) の L_0 を考える。L_0 をどのように最大化するかは、$p_\theta(\mathbf{x}_0|\mathbf{x}_1)$ をどのように指定するかに依存する。そして、この $p_\theta(\mathbf{x}_0|\mathbf{x}_1)$ は、直接的に観測データをモデル化する分布である。例えば、[?] の Section 3.3 を参照されたい。

注意　ここでは、denoising diffusion probabilistic models の変分推論だけを議論している。このモデルがどこから来たのかについては、議論していない。(この点については [?] を参照。)

Contents

Appendix

$$
\begin{aligned}
&\int \exp\left(-\frac{(x-ay)^2}{2s^2}-\frac{(y-bz)^2}{2t^2}\right)dy = \int \exp\left(-\frac{t^2(x-ay)^2+s^2(y-bz)^2}{2s^2t^2}\right)dy \\
&= \int \exp\left(-\frac{(s^2+t^2a^2)y^2-2(s^2bz+t^2ax)y+t^2x^2+s^2b^2z^2}{2s^2t^2}\right)dy \\
&= \exp\left(-\frac{t^2x^2+s^2b^2z^2}{2s^2t^2}\right)\int \exp\left(-\frac{s^2+t^2a^2}{2s^2t^2}\left(y^2-\frac{2(s^2bz+t^2ax)}{s^2+t^2a^2}y\right)\right)dy \\
&= \exp\left(-\frac{t^2x^2+s^2b^2z^2}{2s^2t^2}+\frac{(s^2bz+t^2ax)^2}{2s^2t^2(s^2+t^2a^2)}\right)\int \exp\left(-\frac{s^2+t^2a^2}{2s^2t^2}\left(y-\frac{s^2bz+t^2ax}{s^2+t^2a^2}\right)^2\right)dy \\
&\propto \exp\left(-\frac{s^2t^2x^2+s^4b^2z^2+t^4a^2x^2+s^2t^2a^2b^2z^2-t^4a^2x^2-2s^2t^2abzx-s^4b^2z^2}{2s^2t^2(s^2+t^2a^2)}\right) \\
&= \exp\left(-\frac{x^2-2abzx+a^2b^2z^2}{2(s^2+t^2a^2)}\right) = \exp\left(-\frac{(x-abz)^2}{2(s^2+t^2a^2)}\right) \qquad (34)
\end{aligned}
$$