Tweedle Formula

$$\underset{\text{ruined}}{Y} = \underset{\text{clean}}{X} + \sigma \underset{\text{noise}}{Z}$$

(경험적) 평균, 분산

$$\mathbb{E}[X \mid Y=y] = y + \sigma^2 \nabla_y \log P_Y(y)$$

$$\mathrm{Var}[X \mid Y=y] = \sigma^2 \left(I + \sigma^2 \nabla_y^2 \log P_Y(y) \right)$$

Note. 설득 흐름

$$\mathbb{E}[X \mid Y=y] = y + \underbrace{\sigma^2}_{\text{불확실}} \underbrace{\nabla_y \log P_Y(y)}_{\text{"키가 커요"}}$$

---

① $P_{Y|X}(y|x) = \underbrace{\phi_\sigma (y-x)}_{\text{가우시안}}$

② $\nabla P_Y(y) = \int P_X(x) \left(\nabla_y \phi_\sigma (y-x) \right) dx$

$\downarrow$ 미분

$$= \int P_X(x) \left(-\frac{1}{\sigma^2} (y-x) \phi_\sigma (y-x) \right) dx$$

$\because$ $\phi_s = C \cdot \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right)$

$$\nabla \phi_s = \underbrace{C \cdot \exp(\cdots)}_{= \phi_s} \underbrace{\nabla_u \left(-\frac{\|u\|^2}{2\sigma^2}\right)}_{-\frac{u}{\sigma^2}}$$

③ $$= -\frac{1}{\sigma^2} \left[y P_Y(y) - P_Y(y) \underbrace{\int x P_{X|Y}(x|y) dx}_{= \mathbb{E}_{X|Y}[X \mid Y=y]} \right]$$

$\because$ $P_X(x) P_{Y|X}(y|x) = P_{X|Y}(x|y) P_Y(y)$

④ $$\nabla \log P_Y(y) = -\frac{1}{\sigma^2} \left[y - \mathbb{E}_{X|Y}[X \mid Y=y] \right]$$

$\because$ $\nabla \log P_Y(y) = \frac{\nabla P_Y(y)}{P_Y(y)}$

$Y = X + \sigma Z$ when $Z \sim N(0, I)$

then if $\varepsilon \approx 0$

$p_{X|Y}(x|y)$ is Gaussian

가우시안 복잡

① $$p_{X|Y}(X|Y) = \frac{p_{Y|X}(Y|X) \cdot p_X(X)}{\cancel{p_Y(Y)}_{\text{const}}}$$

② $$p_{Y|X}(Y|X) = \underbrace{\cancel{\frac{1}{(2\pi\sigma^2)^{d/2}}}}_{\text{const}} \cdot \exp\left(-\frac{\|y-x\|^2}{2\sigma^2}\right)$$

---

③ $$p_X(x) \approx p_X(y) + \langle \nabla p_X(y), x-y \rangle + \underbrace{\cancel{O(\|X-Y\|^2)}}_{\text{small}}$$

$$= p_X(y) \langle \nabla \log p_X(y), x-y \rangle$$

$$= \cancel{p_X(y)}\left(1 + \langle \nabla \log p_X(y), x-y \rangle\right)$$

복잡

$$\approx \underbrace{\cancel{p_X(y)}}_{\text{const}} \cdot \exp\left(\langle \nabla \log p_X(y), x-y \rangle\right) \qquad (1+x \approx e^x)$$

가우시안 복잡

③ $$p_{X|Y}(X|Y) \propto \exp\left(-\frac{\|y-x\|^2}{2\sigma^2} + \langle \nabla \log p_X(y), x-y \rangle\right)$$

$$= (\cdots) = -\frac{1}{2\sigma^2}\left\| (x-y) - \sigma^2 \nabla \log p_X(y) \right\|^2 + \text{Const}$$

$$\approx N\left(y + \sigma^2 \nabla \log p_X(y),\ \sigma^2 I\right)$$

DDPM.

(forward) $x_t | x_{t-1} = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_t \sim N(0, I)$

$\sim N(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t) I)$

Note $\alpha_t + \beta_t = 1$

$0 < \beta_t < 1$

$x_2 = \sqrt{\alpha_2}\, x_1 + \sqrt{1-\alpha_2}\, z_2$

$= \sqrt{\alpha_2}\sqrt{\alpha_1}\, x_0 + \sqrt{\alpha_2}\sqrt{1-\alpha_1}\, z_1 + \sqrt{1-\alpha_2}\, z_2$

$= \sqrt{\alpha_2}\sqrt{\alpha_1}\, x_0 + \sqrt{1-\alpha_1\alpha_2}\, \tilde{z}$

$\because$ $X_1 \sim N(0, \sigma_1^2 I)$

$X_2 \sim N(0, \sigma_2^2 I)$

then $X_1 + X_2$

$\sim N(0, (\sigma_1^2 + \sigma_2^2) I)$

---

hence, in general forward pass

$$x_t | x_0 \sim N(\underbrace{\sqrt{\alpha_1 \cdots \alpha_t}}_{= \bar{\alpha}_t}\, x_0, (1 - \underbrace{\alpha_1 \cdots \alpha_t}_{= \bar{\alpha}_t}) I)$$

이것은 Tweedie Formula에서

다루는 그 원리와 같다.