Q1. 給定 $q(x_{1:T}|x_0)=\prod_{t=1}^{T} q(x_t|x_{t-1})$

根據 Markvo 的定理可以得到：$\prod_{t=1}^{T} q(x_t|x_{t-1},x_0)$ —①

對於 difussion 來說，是從 x_T 一路消除 noise 以得到與 x_0 相似的圖片。

由①⇒ $q(x_1|x_0)\prod_{t=2}^{T} q(x_t|x_{t-1},x_0)$

$$= q(x_1|x_0)\prod_{t=2}^{T}\frac{q(x_{t-1}|x_t,x_0)\,q(x_t|x_0)\,\cancel{q(x_0)}}{q(x_{t-1}|x_0)\cdot\cancel{q(x_0)}}$$

$$= q(x_1|x_0)\prod_{t=2}^{T}\frac{q(x_t|x_0)}{q(x_{t-1}|x_0)}\prod_{t=2}^{T} q(x_{t-1}|x_t,x_0)$$

$$= q(x_1|x_0)\left(\frac{\cancel{q(x_2|x_0)}}{q(x_1\cdot x_0)}\cdot\frac{\cancel{q(x_3|x_0)}}{\cancel{q(x_2|x_0)}}\cdot\frac{\cancel{q(x_4|x_0)}}{\cancel{q(x_3|x_0)}}\cdots\frac{q(x_T|x_0)}{\cancel{q(x_{T-1}|x_0)}}\right)\prod_{t=2}^{T} q(x_{t-1}|x_t,x_0)$$

$$= q(x_T|x_0)\cdot\prod_{t=2}^{T} q(x_{t-1}|x_t,x_0)$$

Q2.

$$q(\mathbf{x}_t|\mathbf{x}_0)=\mathcal{N}(\mathbf{x}_t;\sqrt{\bar\alpha_t}\mathbf{x}_0,(1-\bar\alpha_t)\mathbf{I})$$

proof 4

$$x_t=\sqrt{\alpha_t}\,\underline{x_{t-1}}+\sqrt{1-\alpha_t}\,\varepsilon_t\ ,\quad \varepsilon_t\sim N(0,1)$$

$$\Rightarrow \sqrt{\alpha_t}\left(\underline{\sqrt{\alpha_{t-1}}\,x_{t-2}+\sqrt{1-\alpha_{t-1}}\,\varepsilon_{t-1}}\right)+\sqrt{1-\alpha_t}\,\varepsilon_t$$

$$\Rightarrow \sqrt{\alpha_t\cdot\alpha_{t-1}}\,x_{t-2}+\sqrt{\alpha_t(1-\alpha_{t-1})}\,\varepsilon_{t-1}+\sqrt{1-\alpha_t}\,\varepsilon_t$$

$$\Rightarrow \sqrt{\alpha_t\cdot\alpha_{t-1}}\left(\sqrt{\alpha_{t-2}}\,x_{t-3}+\sqrt{1-\alpha_{t-2}}\,\varepsilon_{t-2}\right)+\sqrt{\alpha_t(1-\alpha_{t-1})}\,\varepsilon_{t-1}+\sqrt{1-\alpha_t}\,\varepsilon_t$$

$$\Rightarrow \sqrt{\alpha_t\cdot\alpha_{t-1}\cdot\alpha_{t-2}}\,x_{t-3}+\sqrt{\alpha_t\cdot\alpha_{t-1}(1-\alpha_{t-2})}\,\varepsilon_{t-2}+\sqrt{\alpha_t(1-\alpha_{t-1})}\,\varepsilon_{t-1}+\sqrt{1-\alpha_t}\,\varepsilon_t$$

$$\vdots$$

$$\Rightarrow \sqrt{\alpha_t\cdot\alpha_{t-1}\cdots\alpha_1}\,x_0+\left(\underline{\sqrt{\alpha_t\cdots\alpha_2(1-\alpha_1)}\,\varepsilon_1+\cdots+\sqrt{\alpha_t(1-\alpha_{t-1})}\,\varepsilon_{t-1}+\sqrt{1-\alpha_t}\,\varepsilon_t}\right)$$

獨立的 Gaussian noise $\sim N(0,\bar\alpha_t)$，其中 $\bar\alpha_t=1-(\alpha_t\cdots\alpha_1)$

接續上頁的 total variance，設 $\alpha = 1-\beta$

$$\underline{\alpha_t\alpha_{t-1}\cdots\alpha_1 + \alpha_t\alpha_{t-1}\cdots\alpha_2\beta_1} + \alpha_t\alpha_{t-1}\cdots\alpha_3\beta_2 + \cdots + \alpha_t\beta_{t-1} + \beta_t$$

$$= \underline{(\alpha_t\alpha_{t-1}\cdots\alpha_2)(\alpha_1+\beta_1)} + \alpha_t\alpha_{t-1}\cdots\alpha_3\beta_2 + \cdots + \alpha_t\beta_{t-1} + \beta_t$$

$$= \underline{(\alpha_t\alpha_{t-1}\cdots\alpha_2)\cdot 1 + \alpha_t\alpha_{t-1}\cdots\alpha_3\beta_2} + \cdots + \alpha_t\beta_{t-1} + \beta_t$$

$$= \underline{(\alpha_t\alpha_{t-1}\cdots\alpha_3)(\alpha_2+\beta_2)} + \cdots + \alpha_t\beta_{t-1} + \beta_t$$

$$= (\alpha_t\alpha_{t-1}\cdots\alpha_3)\cdot 1 + \cdots + \alpha + \beta_{t-1} + \beta_t$$

$$\vdots$$

$$= \alpha_t + \beta_t = 1$$

∴ noise variance

$$= 1 - \alpha_t\alpha_{t-1}\cdots\alpha_1$$
$$= 1 - \bar{\alpha}_t$$

$$\therefore\ q(x_t|x_0) = N(x_t; \sqrt{\bar{\alpha}_t}\,x_0, (1-\bar{\alpha}_t)I)$$

proof 6

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

從方程式4中，可得 x_t：

$$x_t = x_0\sqrt{\bar{\alpha}_t} + \sqrt{1-\bar{\alpha}_t}\,\epsilon_t$$

$$\Rightarrow x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t - \sqrt{1-\bar{\alpha}_t}\,\epsilon_t)$$

if $q(x_t|x_{t-1}) \sim N$ 且 $\beta_t = (1-\alpha_t)$ 夠小，則 $q(x_{t-1}|x_t) \sim N$

$$\therefore\ q(x_{t-1}|x_t) = q(x_{t-1}|x_t, x_0)$$
$$= N(x_{t-1}, \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t I)$$

proof 8

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

根據 DDPM 中的方程式5

$$\mathbb{E}_q\Big[\underbrace{D_{\mathrm{KL}}(q(\mathbf{x}_T|\mathbf{x}_0)\,\|\,p(\mathbf{x}_T))}_{L_T} + \sum_{t>1}\underbrace{D_{\mathrm{KL}}(q(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)\,\|\,p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t))}_{L_{t-1}}\underbrace{-\log p_\theta(\mathbf{x}_0|\mathbf{x}_1)}_{L_0}\Big] \quad (5)$$

L_T 是 forward process 可忽略，L_0 是固定

∴ 只須看 L_{t-1} 項，已知 $q(x_{t-1}|x_t,x_0) \sim N$，$p_\theta(x_{t-1}|x_t) \sim N$

可知 $D_{KL}(N(\mu_1,\sigma_1^2)\,\|\,N(\mu_2,\sigma_2^2))$

$$= \int \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \cdot \log\frac{\frac{1}{\sqrt{2\pi}\sigma_1}\cdot e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}}}{\frac{1}{\sqrt{2\pi}\sigma_2}\cdot e^{\frac{-(x-\mu_2)^2}{2\sigma_2^2}}}\,dx$$

$$= \int \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \cdot \left(\log\frac{\sigma_2}{\sigma_1} - \frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(x-\mu_2)^2}{2\sigma_2^2} \right)$$

$$= \int \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \cdot \log\frac{\sigma_2}{\sigma_1} - \int \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{(x-\mu_1)}{2\sigma_1^2} + \int \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{(x-\mu_2)^2}{2\sigma_2^2}$$

$$= \log\frac{\sigma_2}{\sigma_1} - \frac{1}{2} + \frac{\sigma_1^2 + (\mu_1-\mu_2)^2}{2\sigma_2^2}$$

$$\therefore L_{t-1} = D_{KL}\left(q(x_{t-1} \mid x_t, x_0) \,\|\, p_\theta(x_{t-1} \mid x_t)\right)$$

$$= D_{KL}\left(N(x_{t-1}; \tilde{\mu}_t(x_t, x_0), \sigma_t^2 I \,\|\, N(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I)\right)$$

$$= E_q\left[\frac{1}{2\sigma_t^2} \| \tilde{\mu}_t(x_t, x_0) - \mu_\theta(x_t, t) \|^2 \right] + C$$