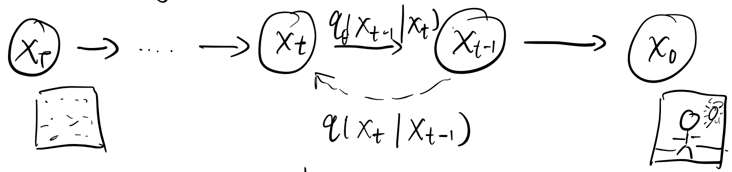
I. Set up denoising diffusion probablistic model (DDPM) (Ho et al, 2020)



1.1 The Forward Noising chain:

$$q(x_{1:T}|x_0) = \prod_{i=1}^{T} q(x_t|x_{t-1}), \quad q(x_t|x_{t-1}) = N(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t)I)$$

we write $1-\alpha_t = \beta_t$, where $\beta_t \in (0,1)$

equivalently, one forward step is

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

1.2 Derive the closed form $q(x_t|x_0)$:

Define $\bar{\alpha}_t := \prod_{s=1}^{t} \alpha_s$, since

$$\begin{cases} x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\,\epsilon_t \\ x_{t-1} = \sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_{t-1}}\,\epsilon_{t-1} \\ \vdots \\ x_1 = \sqrt{\alpha_1}\, x_0 + \sqrt{1-\alpha_1}\,\epsilon_1 \end{cases}$$

$$x_t = \sqrt{\alpha_t}\sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{\alpha_t}\sqrt{1-\alpha_{t-1}}\,\epsilon_{t-1} + \sqrt{1-\alpha_t}\,\epsilon_t$$

$$= \sqrt{\alpha_t}\sqrt{\alpha_{t-1}}\sqrt{\alpha_{t-2}}\, x_{t-3} + \sqrt{\alpha_t}\sqrt{\alpha_{t-1}}\sqrt{1-\alpha_{t-2}}\,\epsilon_{t-2} + \sqrt{\alpha_t}\sqrt{1-\alpha_{t-1}}\,\epsilon_{t-1} + \sqrt{1-\alpha_t}\,\epsilon_t$$

$$\vdots$$

$$x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sum_{s=1}^{t} \sqrt{\underbrace{(1-\alpha_s)}_{\text{constant}} \underbrace{\prod_{r=s+1}^{t} \alpha_r}_{\text{const.}}}\ \epsilon_s \rightarrow \text{Gaussian} \quad (\text{if } r = t+1 \ \ \prod_{r=t+1}^{t} \alpha_r = 1)$$

an i.i.d. g. $\quad g \sim N(0, \sum_{s=1}^{t}(1-\alpha_s)\prod_{r=s+1}^{t}\alpha_r I)$

Note that:

$$\sum_{s=1}^{t}(1-\alpha_s)\prod_{r=s+1}^{t}\alpha_r = \sum_{s=1}^{t}\left(\prod_{r=s+1}^{t}\alpha_r - \alpha_s\prod_{r=s+1}^{t}\alpha_r\right)$$

$$= \sum_{s=1}^{t}\left(\prod_{r=s+1}^{t}\alpha_r - \prod_{r=s}^{t}\alpha_r\right) = \prod_{r=2}^{t}\alpha_r - \prod_{r=1}^{t}\alpha_r + \prod_{r=3}^{t}\alpha_r - \prod_{r=2}^{t}\alpha_r + \dots + \prod_{r=t}^{t}\alpha_r - \prod_{r=t-1}^{t}\alpha_r$$

$$= 1 - \prod_{r=1}^{t}\alpha_r = 1 - \bar{\alpha}_t.$$

$$\Rightarrow x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\,\epsilon, \quad \epsilon \sim N(0, I)$$

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

Why important? For training, although forward process is defined as T-step MC, we can still jump directly from x_0 to x_t.

1.3 The reverse posterior $q(x_{t-1}|x_t, x_0)$:

Given $q(x_t|x_{t-1})$ and $q(x_{t-1}|x_0)$ are Gaussian. $q(x_{t-1}|x_0, x_t)$ is also Gaussian.

$$q(x_{t-1}|x_0, x_t) \propto q(x_t|x_{t-1}) \cdot q(x_{t-1}|x_0) \propto \exp(m_1 + m_2).$$

$$q(x_t|x_{t-1}) \propto \underbrace{\exp\left(-\frac{1}{2(1-\alpha_t)}\|x_t - \sqrt{\alpha_t}\, x_{t-1}\|^2\right)}_{m_1}$$

$$q(x_{t-1}|x_0) = N(\sqrt{\bar\alpha_{t-1}}\, x_0, (1-\bar\alpha_{t-1})I).$$

$$\propto \exp\left(-\frac{1}{2(1-\bar\alpha_{t-1})}\|x_{t-1}-\sqrt{\bar\alpha_{t-1}}\, x_0\|^2\right) \quad m_2.$$

$$m_1+m_2 = -\frac{1}{2(1-\alpha_t)}\|x_t-\sqrt{\alpha_t}\, x_{t-1}\|^2 - \frac{1}{2(1-\bar\alpha_{t-1})}\|x_{t-1}-\sqrt{\bar\alpha_{t-1}}\, x_0\|^2$$

open squares ⇓

$$= -\frac{1}{2}\left[\left(\frac{\alpha_t}{1-\alpha_t}+\frac{1}{1-\bar\alpha_{t-1}}\right)\|x_{t-1}\|^2 - 2\left(\frac{\sqrt{\alpha_t}}{1-\alpha_t}x_t + \frac{\sqrt{\bar\alpha_{t-1}}}{1-\bar\alpha_{t-1}}x_0\right)^T x_{t-1}\right] + \text{const.}$$

Completing the square:

$$q(x_{t-1}|x_t, x_0) = N(\tilde\mu_t(x_t, x_0), \tilde\beta_t I), \quad \text{("True reverse")}$$

where $\tilde\beta_t = \frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}(1-\alpha_t) = \frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\beta_t$

and

$$\tilde\mu_t(x_t, x_0) = \frac{\sqrt{\alpha_t}(1-\bar\alpha_{t-1})}{1-\bar\alpha_t}x_t + \frac{\sqrt{\bar\alpha_{t-1}}\beta_t}{1-\bar\alpha_t}x_0$$

⇒ if I knew x_0, reverse posterior can be derived explicitly,

But during generation we don't know x_0 !!!

Recall $x_t = \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\,\epsilon$, so $x_0 = \frac{1}{\sqrt{\bar\alpha_t}}(x_t - \sqrt{1-\bar\alpha_t}\,\epsilon)$

If the model can predict ϵ, given x_t, t, (rewrite as $\epsilon_\theta(x_t, t)$) $\quad \frac{\sqrt{\bar\alpha_{t-1}}}{\sqrt{\bar\alpha_t}} = \frac{1}{\sqrt{\alpha_t}}$

Then we have $\hat{x}_0(x_t, t) = \frac{1}{\sqrt{\bar\alpha_t}}(x_t - \sqrt{1-\bar\alpha_t}\,\epsilon_\theta(x_t, t))$.

Then

$$\mu_\theta(x_t, x_0) = \frac{\beta_t}{(1-\bar\alpha_t)}\frac{1}{\sqrt{\alpha_t}}(x_t - \sqrt{1-\bar\alpha_t}\,\epsilon_\theta(x_t, t)) + \frac{\sqrt{\alpha_t}(1-\bar\alpha_{t-1})}{1-\bar\alpha_t}x_t$$

$$= \frac{\beta_t + \alpha_t - \bar\alpha_t}{(1-\bar\alpha_t)\sqrt{\alpha_t}}x_t - \frac{\beta_t\sqrt{1-\bar\alpha_t}\,\epsilon_\theta(x_t, t)}{(1-\bar\alpha_t)\sqrt{\alpha_t}}$$

$$= \frac{1-\alpha_t+\alpha_t-\bar\alpha_t}{(1-\bar\alpha_t)\sqrt{\alpha_t}}x_t - \frac{\beta_t\sqrt{1-\bar\alpha_t}\,\epsilon_\theta(x_t, t)}{(1-\bar\alpha_t)\sqrt{\alpha_t}}$$

$$= \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t\sqrt{1-\bar\alpha_t}}{1-\bar\alpha_t}\epsilon_\theta(x_t, t)\right)$$

1.4: Parameterize the model:

$$p_\theta(x_{0:T}) = p(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t), \quad p(x_T) \sim N(0, I)$$

$$p_\theta(x_{t-1}|x_t) = N(\underbrace{\mu_\theta(x_t, t)}_{\downarrow\ \text{NN}}, \underbrace{\sigma_t^2}_{\downarrow\ \text{hyperpara}} I)$$

1.5 Loss:

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

Recall: $\begin{cases} q(x_{t-1} | x_t, x_0) = N(\tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t I) \\ p_\theta(x_{t-1} | x_t) = N(\mu_\theta(x_t, t), \sigma_t^2 I) \end{cases}$

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

Then from 1.3.

$$\tilde{\mu}_t(x_t, x_0) - \mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}}\left[x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\varepsilon - (x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\varepsilon_\theta(x_t, t))\right]$$

$$= \frac{\beta_t}{\sqrt{\alpha_t}\sqrt{1-\bar{\alpha}_t}}(\varepsilon_\theta(x_t, t) - \varepsilon)$$

Therefore:

$$\|\tilde{\mu}_t(x_t, x_0) - \mu_\theta(x_t, t)\|^2 = \frac{\beta_t^2}{\alpha_t(1-\bar{\alpha}_t)} \|\varepsilon - \varepsilon_\theta(x_t, t)\|^2$$

$$\Rightarrow L_{t-1} = E_{x_0, \varepsilon}\left[\frac{\beta_t^2}{2\sigma_t^2 \alpha_t (1-\bar{\alpha}_t)} \|\varepsilon - \varepsilon_\theta(x_t, t)\|^2\right] + C_t$$

$$L_{simple} = E_{x_0, \varepsilon, t}\left[\|\varepsilon - \varepsilon_\theta(x_t, t)\|^2\right]$$