20/01/25

Computer Vision

- Recap
  - Diffusion Process → fwd and bwd trajectories have the same dist func form
  - Maximizing log likelihood ✓
  - Markov chain assumption ✓
  - Conditional dist Gaussian → closed form expression for bound
- Lower bound on the expected log likelihood
- Recap:

$$\mathcal{L}(\theta) = \int q(x^{(0)}) \cdot \log p_\theta(x^{(0)})\, dx^{(0)} \quad -①$$

$$\theta^* = \underset{\theta \in \Theta}{\arg\max}\ \mathcal{L}(\theta) \quad -②$$

$$p_\theta(x^{(0)}) = \int p_\theta(x^{(0 \dots T)}) \cdot dx^{(1 \dots T)}$$

$$= \int p_\theta(x^{(T)}) \cdot \prod_{t=1}^{T} p_\theta(x^{(t-1)} | x^{(t)})\, dx^{(1 \dots T)}$$

$$p_\theta(x^{(0)}) = \int p_\theta(x^{(T)}) \cdot q(x^{(1 \dots T)} | x^{(0)}) \frac{\prod_{t=1}^{T} p_\theta(x^{(t-1)} | x^{(t)})}{\prod_{t=1}^{T} q(x^{(t)} | x^{(t-1)})}\, dx^{(1 \dots T)} \quad -③$$

Plug ③ into ①

$$\mathcal{L}(\theta) = \int q(x^{(0)}) \cdot \log \Big[\underbrace{\int p_\theta(x^{(T)}) \cdot q(x^{(1 \dots T)} | x^{(0)}) \frac{\prod p_\theta(x^{(t-1)} | x^{(t)})}{\prod q(x^{(t)} | x^{(t-1)})}\, dx^{(1 \dots T)}}_{p_\theta(x^{(0)})} \Big]\, dx^{(0)}$$

We will apply the Jensen's inequality. For a concave function $\varphi(x)$, $\varphi(E[x]) \geq E[\varphi(x)]$

$\mathcal{L}(\theta) \geq K$, where

$$K = \int q(x^{(0)}) \cdot q(x^{(1 \cdots T)} \mid x^{(0)}) \log p_\theta(x^{(T)}) \cdot \frac{\prod p_\theta(x^{(t-1)} \mid x^{(t)})}{\prod q(x^{(t)} \mid x^{(t-1)})} \, dx^{(0 \cdots T)}$$

$$= \int q(x^{(0 \cdots T)}) \cdot \log p_\theta(x^{(T)}) \cdot \frac{\prod p_\theta(x^{(t-1)} \mid x^{(t)})}{\prod q(x^{(t)} \mid x^{(t-1)})} \, dx^{(0 \cdots T)} \quad \text{— (4)}$$

Since the forward trajectory is a Markov chain,

$$q(x^{(t)} \mid x^{(t-1)}) = q(x^{(t)} \mid x^{(t-1)}, x^{(0)})$$

Since $x^{(0)}$ is the image RV

$$= q(x^{(t-1)} \mid x^{(t)}, x^{(0)}) \cdot \frac{q(x^{(t)} \mid x^{(0)})}{q(x^{(t-1)} \mid x^{(0)})} \quad \text{— (5)}$$

Plug (5) into (4)

$$K = \int q(x^{(0 \cdots T)}) \cdot \log p_\theta(x^{(T)}) \cdot \frac{\prod p_\theta(x^{(t-1)} \mid x^{(t)}) \cdot q(x^{(t-1)} \mid x^{(0)})}{\prod q(x^{t-1} \mid x^{(t)}, x^{(0)}) \cdot q(x^{(t)} \mid x^{(0)})} \, dx^{(0 \cdots T)} \quad \text{— (6)}$$

Check: we want to express ⑥ as a sum of KL divergence terms and entropy terms.

$$K = \int q(x^{(0\cdots T)}) \cdot \log p_\theta(x^{(T)})\, dx^{(0\cdots T)} +$$

$$\int q(x^{(0\cdots T)}) \cdot \log\left[\frac{\prod p_\theta(x^{(t-1)}|x^{(t)})}{\prod q(x^{(t-1)}|x^{(t)},x^{(0)})}\frac{q(x^{(t-1)}|x^{(0)})}{q(x^{(t)}|x^{(0)})}\right] dx^{(0\cdots T)}$$