24/01/25

Computer Vision

- Recap:
  - Diffusion process - forward and backward trajectories
  - Markov chains: $q(x^{(0\cdots T)}) = q(x^{(0)}) \cdot \prod_{t=1}^{T} q(x^{(t)} \mid x^{(t-1)})$

    $$p_\theta(x^{(0\cdots T)}) = p_\theta(x^{(T)}) \cdot \prod_{t=1}^{T} p_\theta(x^{(t-1)} \mid x^{(t)})$$

  - Marginal: $p_\theta(x^{(0)}) = \int p_\theta(x^{(0\cdots T)}) \cdot dx^{(1\cdots T)}$
  - Rewriting: $$p_\theta(x^{(0)}) = \int q(x^{(1\cdots T)} \mid x^{(0)}) \cdot \frac{\overbrace{\prod_{t=1}^{T} p_\theta(x^{(t-1)} \mid x^{(t)})}^{\text{Reverse}}}{\underbrace{\prod_{t=1}^{T} q(x^{(t)} \mid x^{(t-1)})}_{\text{Forward}}} \cdot p_\theta(x^{(T)}) \cdot dx^{(1\cdots T)}$$
  - Problem: $\max L(\theta) = \max \int q(x^{(0)}) \cdot \log p_\theta(x^{(0)}) \cdot dx^{(0)}$
  - Lower bound: $L(\theta) \geq K(\theta)$ (Jensen's inequality)

    $$K(\theta) = \int q(x^{(0\cdots T)}) \cdot \log\left[p_\theta(x^{(0)}) \cdot \frac{\prod_{t=1}^{T} p_\theta(x^{(t-1)} \mid x^{(t)})}{\prod_{t=1}^{T} q(x^{(t)} \mid x^{(t-1)})}\right] \cdot dx^{(0\cdots T)} \quad \text{— ①}$$

  - Using Markov property: $$q(x^{(t)} \mid x^{(t-1)}) = q(x^{(t)} \mid x^{(t-1)}, \overbrace{x^{(0)}}^{\text{Image input}})$$

    $$= \frac{q(x^{(t-1)} \mid x^{(t)}, x^{(0)}) \cdot q(x^{(t)} \mid x^{(0)})}{q(x^{(t-1)} \mid x^{(0)})} \quad \text{— ②}$$

  - Plugging ② in ①, we get:

    $$K(\theta) = \int q(x^{(0\cdots T)}) \left[\log p_\theta(x^{(T)}) + \log \prod_{t=1}^{T} \frac{p_\theta(x^{(t-1)} \mid x^{(t)})}{q(x^{(t-1)} \mid x^{(t)}, x^{(0)})} \cdot \frac{q(x^{(t-1)} \mid x^{(0)})}{q(x^{(t)} \mid x^{(0)})}\right] dx^{(0\cdots T)} \quad \text{— ③}$$

- Recall: $KL\left(\mathcal{N}(x;\mu_0,\Sigma_0)\,\|\,\mathcal{N}(x;\mu_1,\Sigma_1)\right)$ $\quad (x\in\mathbb{R}^d)$

$$= \frac{1}{2}\left(tr\left(\Sigma_1^{-1}\Sigma_0\right) + \underbrace{(\mu_1-\mu_0)^T\cdot\Sigma_1^{-1}\cdot(\mu_1-\mu_0)}_{\text{key to the optimization}} - d + \log\left(\frac{\det\Sigma_1}{\det\Sigma_0}\right)\right) \quad -(*)$$

$$\mathcal{N}(x;\mu,\Sigma) = \frac{1}{(2\pi)^{d/2}(\det\Sigma)^{1/2}}\exp\left[-\frac{1}{2}\cdot(x-\mu)^T\cdot\Sigma^{-1}\cdot(x-\mu)\right]$$

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

$$K(\theta) = -\sum_{t=2}^{T}\int q(x^{(0)},x^{(t)})\,KL\left(\underbrace{q(x^{(t-1)}\mid x^{(t)},x^{(0)})}\,\|\,p_\theta(x^{(t-1)}\mid x^{(t)})\right)dx^{(0)}dx^{(t)}$$
$$+ H_q(x^{(T)}\mid x^{(0)}) - H_q(x^{(1)}\mid x^{(0)}) - H_p(x^{(T)}) \quad -④$$

HW: Please verify the expression in ④

Note: All the terms in ④ can be evaluated exactly under the following conditions:

- $\underbrace{q(x^{(t)}\mid x^{(t-1)})}_{\beta_t:\ \text{deterministic hyperparameter}} \sim \mathcal{N}\left(x^{(t)};\sqrt{1-\beta_t}\cdot x^{(t-1)},\beta_t I\right) \quad -⑤$

- $p_\theta(x^{(t-1)}\mid x^{(t)}) \sim \mathcal{N}\left(x^{(t-1)}; f_\mu(x^{(t)},t;\theta_\mu), f_\Sigma(x^{(t)},t;\theta_\Sigma)\right) \quad -⑥$

• $q(x^{(t-1)} \mid x^{(t)}, x^{(0)}) \sim \mathcal{N}(x^{(t-1)}; \tilde{\mu}(x^{(t)}, x^{(0)}), \tilde{\beta}_t I)$ — (7)

Key term: $E_q[\| \tilde{\mu}(\quad) - f_\theta(\quad) \|^2]$