DDPM Training Objective Proof

Aug. 16, 2023

Problem 1. Please provide a detailed mathematical proof for the training objective of the Denoising Diffusion Probabilistic Model (DDPM), which aims to maximize the log-likelihood of data $\log p(x_0; \theta)$ where θ represents the model parameters. Your proof should demonstrate the derivation started from

$$\mathbb{E}_q \left[-\log \frac{p_\theta(\boldsymbol{x}_{0:T})}{q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)} \right]$$

to the training objective as following

$$\mathbb{E}_q \left[-\log p_\theta(\boldsymbol{x}_0|\boldsymbol{x}_1) + \sum_{t=2}^{T} D_{\mathrm{KL}}(q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0) \,||\, p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)) + D_{\mathrm{KL}}(q(\boldsymbol{x}_T|\boldsymbol{x}_0) \,||\, p(\boldsymbol{x}_T)) \right].$$

Hint: Given $x_0, x_1, ..., x_T$, which form a first order Markov chain. Then we have

$$q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0) = \prod_{t=1}^{T} q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \prod_{t=1}^{T} q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{x}_0).$$