Preliminary Research Ideas (2): Combining Markov Chain Monte Carlo and Generative Modeling

Professor Yida Xu
Machine Learning Knowledge Disseminator

Introduction

When executing Markov Chain Monte Carlo (MCMC) methods, such as Gibbs sampling, we typically start with a known parametric form of the target distribution $p_\theta(x)$ (which is usually defined up to a normalizing constant). The primary objective is to generate samples from this specific distribution.

Conversely, generative modeling—particularly in the context of diffusion models—can be conceptualized as an inverse problem. Here, we do not have explicit access to the true underlying distribution $p_{data}(x)$. Instead, we possess a set of empirical samples (i.e., training data), and the goal is to learn a model $p_\theta(x)$ to approximate $p_{data}(x)$, or at least to learn a mechanism for generating new samples that resemble the training data.

This raises an interesting question: Can these two paradigms work synergistically? For instance, could we train a diffusion model to partially approximate the gradients or energy of the data distribution, and subsequently utilize MCMC techniques (such as Langevin dynamics) to refine the generation process, thereby ensuring that the newly generated samples more accurately match the statistical properties of the original training data?