A Practical Diffusion Path for Sampling

Omar Chehab, Anna Korba

How can we efficiently sample from a target distribution π knowing its score?

Background

Sampling process: Annealed Langevin Dynamics

$$x_{k+1} = x_k + h_k \nabla \log \mu_k(x_k) + \sqrt{2h_k}\epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, I)$$

Prescribed path: Convolve target with a Gaussian, using increasing $\lambda_k \in [0,1]$.

$$\mu_k(x) = \frac{1}{\sqrt{\lambda_k}} \pi\left(\frac{x}{\sqrt{\lambda_k}}\right) * \frac{1}{\sqrt{1-\lambda_k}} \mathcal{N}\left(\frac{x}{\sqrt{1-\lambda_k}}; 0, I\right)$$

Estimated score: $\nabla \log \mu_k(x) = \dfrac{\sqrt{\lambda_k}}{1-\lambda_k} \mathbb{E}_{y \sim m_k}\left[y - \dfrac{1}{\sqrt{\lambda_k}} x\right]$,

which requires drawing samples from: $m_k(y \mid x) \propto \pi(y) \times \mathcal{N}\left(y; \dfrac{1}{\sqrt{\lambda_k}} x, \left(\dfrac{1}{\lambda_k} - 1\right) I\right)$

- ❌ Sampling from $m_k(y \mid x)$ is **hard** (non log-concave)
- ❌ Sampling from $m_k(y \mid x)$ is **frequent** (new routine for each x)
- ❌ Estimation error of the score is **big** (exponential in dimension)

Our solution

Choose a Dirac proposal to simplify the convolution.

Explicit score: $\nabla \log \mu_k(x) = \dfrac{1}{\sqrt{\lambda_k}} \nabla \log \pi\left(\dfrac{x}{\sqrt{\lambda_k}}\right)$