Sampling via Nonlinear Diffusion Equations

Women in OT Workshop

Claire Murphy

UC Santa Barbara

April 18, 2024

Motivation

Setup: Let $\tilde{\rho}$ be a probability measure on Euclidean space $\mathbb{R}^d$.

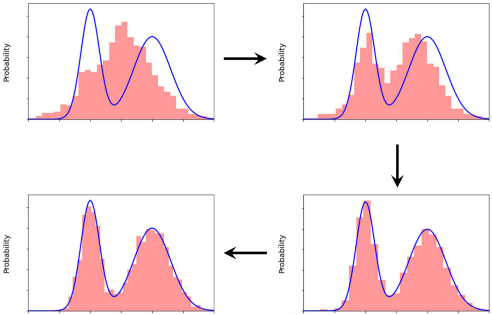
Goal: We seek $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ such that the empirical measure $\frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ converges to $\tilde{\rho}$ as $n \to \infty$.

Motivation

Setup: Let $\tilde{\rho}$ be a probability measure on Euclidean space $\mathbb{R}^d$.

Goal: We seek $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ such that the empirical measure $\frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ converges to $\tilde{\rho}$ as $n \to \infty$.

- Our definition of “convergence" depends on the context of the problem. For example, we may define convergence in terms of the 2-Wasserstein metric.



Source: Science Direct, 2021

Classical Approach: Langevin Dynamics

Assumption: The target measure $\tilde{\rho}$ is strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function $V : \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$.

Classical Approach: Langevin Dynamics

Assumption: The target measure $\tilde{\rho}$ is strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function $V : \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$.

For any initialization $\{x_{i,0}\}_{i=1}^n$, evolving particles by the stochastic differential equation

$$\begin{cases} dx_i(t) = \nabla \log(\tilde{\rho}(x_i))dt + dW_i \\ x_i(0) = x_{i,0}. \end{cases}$$

ensures that $\lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{x_i(t)} = \tilde{\rho}$.

Classical Approach: Langevin Dynamics

Assumption: The target measure $\tilde{\rho}$ is strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function $V : \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$.

For any initialization $\{x_{i,0}\}_{i=1}^n$, evolving particles by the stochastic differential equation

$$\begin{cases} dx_i(t) = \nabla \log(\tilde{\rho}(x_i))dt + dW_i \\ x_i(0) = x_{i,0}. \end{cases}$$

ensures that $\lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i(t)} = \tilde{\rho}$.

Remark (Continuum Perspective)

At time t, the particles approximate $\rho(t,x)$, the solution to the **Fokker-Planck equation**:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \nabla \log(\tilde{\rho})) = \Delta \rho \quad t \geq 0 \\ \rho(0,x) = \rho_0(x). \end{cases}$$

$\rho(t,x)$ converges to $\tilde{\rho}$ as $t \to \infty$.

The Nonlinear Diffusion Equation

A drawback of Langevin dynamics is that the target measure $\tilde{\rho}$ must be strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function V.

The Nonlinear Diffusion Equation

A drawback of Langevin dynamics is that the target measure $\tilde{\rho}$ must be strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function V.

A new approach allows us to consider target measures of the form

$$\tilde{\rho} = ((f')^{-1}(Z - V(x)))_+ dx,$$

where

- Z is a normalizing constant.
- $V : \mathbb{R}^d \to \mathbb{R}$ and $f : [0, \infty) \to \mathbb{R}$ are smooth.
- V is λ-convex for some $\lambda > 0$.
- f is convex and $s \mapsto s^d f(s^{-d})$ is convex and nonincreasing on $(0, \infty)$.

The Nonlinear Diffusion Equation

A drawback of Langevin dynamics is that the target measure $\tilde{\rho}$ must be strongly log-concave, i.e. $\tilde{\rho} = e^{-V(x)}dx$ for a λ-convex function V.

A new approach allows us to consider target measures of the form

$$\tilde{\rho} = ((f')^{-1}(Z - V(x)))_+ dx,$$

where

- Z is a normalizing constant.
- $V : \mathbb{R}^d \to \mathbb{R}$ and $f : [0, \infty) \to \mathbb{R}$ are smooth.
- V is λ-convex for some $\lambda > 0$.
- f is convex and $s \mapsto s^d f(s^{-d})$ is convex and nonincreasing on $(0, \infty)$.

Key idea: If $\rho(t, x)$ is a solution to the **Generalized Fokker-Planck equation**:

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla V) = \nabla \cdot (\rho \nabla f'(\rho)) & t \geq 0 \\ \rho(0, x) = \rho_0(x), \end{cases}$$

then $\rho(t, x)$ still converges to $\tilde{\rho}$ as $t \to \infty$.

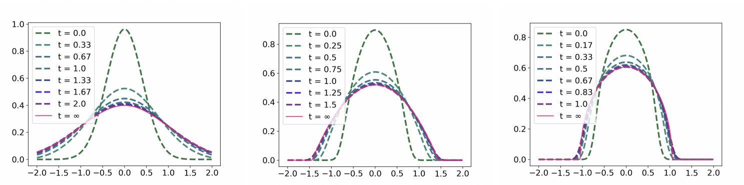
Sampling via Nonlinear Diffusion Equations

Goal: Develop a stochastic particle method to approximate $\rho(t, x)$, the solution to the Generalized Fokker-Planck equation.

Sampling via Nonlinear Diffusion Equations

Goal: Develop a stochastic particle method to approximate $\rho(t, x)$, the solution to the Generalized Fokker-Planck equation.

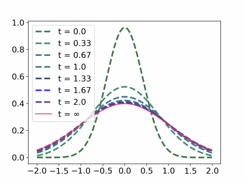
Preliminary results:

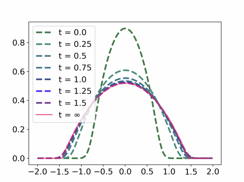


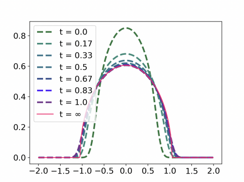
Sampling via Nonlinear Diffusion Equations

Goal: Develop a stochastic particle method to approximate $\rho(t, x)$, the solution to the Generalized Fokker-Planck equation.

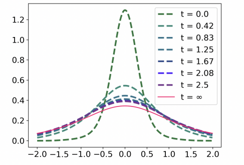
Preliminary results:







Trouble case:



References

- Craig, Katy, Karthik Elamvazhuthi, Matt Haberland, and Olga Turanova. "A blob method for inhomogeneous diffusion with applications to multi-agent control and sampling." Mathematics of Computation (2023).
- Craig, Katy, Matt Jacobs, and Olga Turanova. "A blob method for general nonlinear diffusion." In preparation.
- Jin, Shi, Lei Li, and Jian-Guo Liu. "Random batch methods (RBM) for interacting particle systems." Journal of Computational Physics (2020).

A Particle Method for Generalized Fokker-Planck Equation

For simplicity, assume that f is of the form

$$f(x) = \begin{cases} \frac{x^m}{m-1} & m > 1 \\ x\log(x) & m = 1. \end{cases}$$

This function is associated with the diffusion equation.
Then our particles flow according to the following ODE:

$$\begin{cases} \partial_t x_i(t) = -\nabla V(x_i(t)) - \sum_{j=1}^{n} \left(\left(\sum_{k=1}^{n} (\nabla\varphi(x_k - x_j) m_k \right)^{m-2} + \left(\sum_{k=1}^{n} \varphi(x_i - x_k) m_k \right)^{m-2} \right) \\ \qquad \nabla\varphi(x_j - x_i) m_j. \\ x_i(0) = x_i^0. \end{cases}$$

Here, $\{m_i\}_{i=1}^n$ are defined by $m_i = \frac{\rho_0(x_i)}{n^d}$. φ is a mollifier.