Score-Based Diffusion Generative Models

Jiayu Zhai

April 1, 2025

Table of Contents

Introduction

- **Likelihood-based models**: Directly learn probability density functions (e.g., autoregressive models, normalizing flow models, VAEs).
- **Implicit generative models**: Represent distributions implicitly via sampling processes (e.g., GANs).
- **Limitations**:
  - Likelihood-based models require restrictive architectures or approximations.
  - Implicit models often require unstable adversarial training.
- **Score-based models**: Model the gradient of the log probability density function (score function).

Introduction

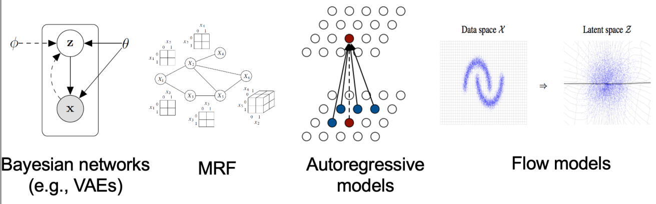

Figure: likelihood based models

Introduction

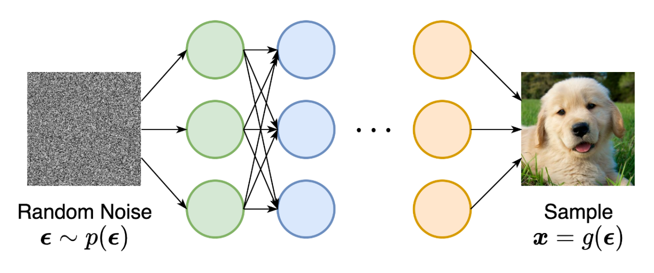

Figure: implicit models

The Score Function and Score Matching

- **Stochastic differential equation**: $\mathbf{x}_{k+1} = \mathbf{x}_k + \sqrt{2\epsilon} \cdot \mathbf{z}_k$
- **Score function**: $\nabla_{\mathbf{x}} \ln p(\mathbf{x})$.
- **Score-based models**: Directly model the score function without worrying about normalizing constants.
- **Score matching**: Train score-based models by minimizing Fisher divergence without adversarial optimization.

Langevin dynamics

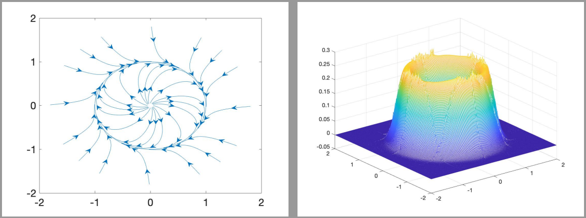

Figure: Stochastic Differential Equations

$\{$

Langevin dynamics

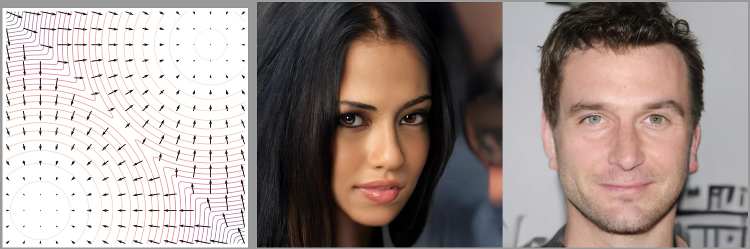

Figure: Score-based diffusion for image

Langevin Dynamics

- Use Langevin dynamics to sample from distributions using only their score functions.
- Iterative procedure:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \epsilon \cdot \nabla_{\mathbf{x}} \ln p(\mathbf{x}_k) + \sqrt{2\epsilon} \cdot \mathbf{z}_k \tag{1}$$

- $\mathbf{z}_k$ is Gaussian noise.

Langevin Dynamics and Score-based Generative Modeling

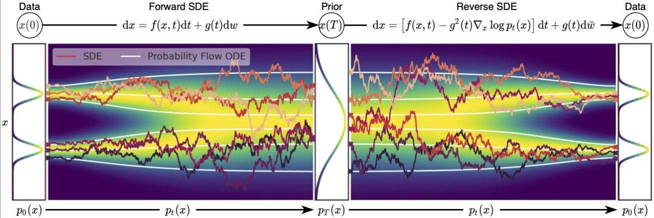

Figure: Caption

Score-based Generative Modeling

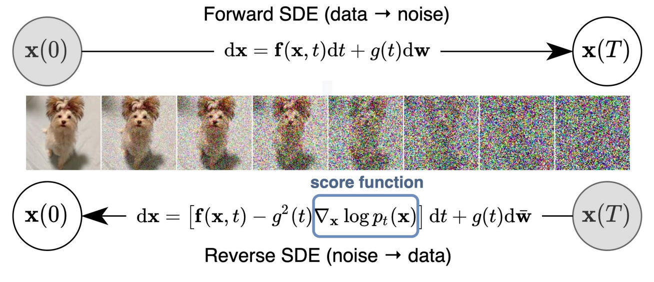

Figure: Caption

Score-based Generative Modeling

The **score function** $\nabla_{\mathbf{x}} \ln p(\mathbf{x})$ is the key for the inverse generative process, but it is in general impossible.
Train a neural network $s_\theta(x)$ to approximate the score function by minimizing

$$\min_{\theta} \| s_\theta(x) - \nabla_{\mathbf{x}} \ln p(\mathbf{x}) \|_{L^2(p)}$$

Score-based Generative Modeling

$$
\begin{aligned}
&\|s_\theta(\mathbf{x}) - \nabla_\mathbf{x} \ln p(\mathbf{x})\|_{L^2(p)} \\
&= \int s_\theta(x) - \nabla_\mathbf{x} \ln p(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \\
&= \int s_\theta^2(x) p(\mathbf{x}) - 2 s_\theta(x) p(\mathbf{x}) \nabla_\mathbf{x} \ln p(\mathbf{x}) + (\nabla_\mathbf{x} \ln p(\mathbf{x}))^2 p(\mathbf{x}) d\mathbf{x} \\
&= \int s_\theta^2(x) p(\mathbf{x}) - 2 s_\theta(x) \nabla_\mathbf{x} p(\mathbf{x}) d\mathbf{x} + C_{\text{data}} \\
&= \int s_\theta^2(x) p(\mathbf{x}) + 2 \nabla_\mathbf{x} s_\theta(x) p(\mathbf{x}) d\mathbf{x} + C_{\text{data}} \\
&= \mathbb{E}_p[s_\theta^2(x) + 2 \nabla_\mathbf{x} s_\theta(x)] + C_{\text{data}}
\end{aligned}
$$

Score-based Generative Modeling

- **Challenge**: Score estimation is inaccurate in low-density regions.
- **Reason**: Fisher divergence minimizes differences weighted by $p(\mathbf{x})$, ignoring low-density regions.
- **Impact**: Poor sample quality when starting from low-density regions.

Score-based Generative Modeling with Multiple Noise Perturbations

- **Solution**: Perturb data with multiple noise scales and train noise-conditional score-based models (NCSN).
- **Advantages**:
  - Improved score estimation accuracy.
  - High-quality sample generation.

Connection to Diffusion Models

- Score-based models and diffusion models are closely related.
- **Key similarities**:
  - Both perturb data with multiple noise scales.
  - Unified framework through stochastic differential equations.

Challenges and Future Directions

- **Challenges**:
  - Slow sampling speed.
  - Difficulty with discrete data.
- **Future directions**:
  - Improve sampling efficiency.
  - Extend to discrete data distributions.

Thank You!

Thank You!