Deep Generative Models

Lecture 14

Roman Isachenko

Moscow Institute of Physics and Technology
Yandex School of Data Analysis

2024, Autumn

Outline

1. Conditional flow matching
   - Conical gaussian paths (continued)
   - Link with score-based models
   - Linear interpolation
   - Rectified flows
2. Latent models
   - Score-based models
   - Autoregressive models
3. The worst course overview

Recap of previous lecture

Discrete-in-time objective

$$\mathbb{E}_{\pi(\mathbf{x}_0)}\mathbb{E}_{t\sim U\{1,T\}}\mathbb{E}_{q(\mathbf{x}_t|\mathbf{x}_0)}\left\|\mathbf{s}_{\boldsymbol{\theta},t}(\mathbf{x}_t) - \nabla_{\mathbf{x}_t}\log q(\mathbf{x}_t|\mathbf{x}_0)\right\|_2^2$$

Continuous-in-time objective

$$\mathbb{E}_{\pi(\mathbf{x}(0))}\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{q(\mathbf{x}(t)|\mathbf{x}(0))}\left\|\mathbf{s}_{\boldsymbol{\theta}}(\mathbf{x}(t), t) - \nabla_{\mathbf{x}(t)}\log q(\mathbf{x}(t)|\mathbf{x}(0))\right\|_2^2$$

NCSN

$$q(\mathbf{x}(t)|\mathbf{x}(0)) = \mathcal{N}\left(\mathbf{x}(0), \left[\sigma^2(t) - \sigma^2(0)\right]\cdot\mathbf{I}\right)$$

DDPM

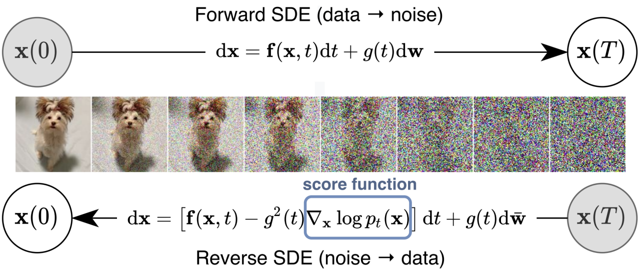
$$q(\mathbf{x}(t)|\mathbf{x}(0)) = \mathcal{N}\left(\mathbf{x}(0)e^{-\frac{1}{2}\int_0^t \beta(s)ds}, \left(1 - e^{-\int_0^t \beta(s)ds}\right)\cdot\mathbf{I}\right)$$

Song Y., et al. Score-Based Generative Modeling through Stochastic Differential Equations, 2020

Recap of previous lecture

Sampling

Solve reverse SDE using numerical solvers (SDESolve).



- Discretization of the reverse SDE gives us the ancestral sampling.
- If we use probability flow instead of SDE than the reverse ODE gives us the DDIM sampling.

Song Y., et al. Score-Based Generative Modeling through Stochastic Differential Equations, 2020

Recap of previous lecture

Let consider ODE dynamic $\mathbf{x}(t)$ in time interval $t \in [0, 1]$ with $\mathbf{x}_0 \sim p_0(\mathbf{x}) = p(\mathbf{x})$, $\mathbf{x}_1 \sim p_1(\mathbf{x}) = \pi(\mathbf{x})$.

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t), \quad \text{with initial condition } \mathbf{x}(0) = \mathbf{x}_0.$$

KFP theorem (continuity equation)

$$\frac{\partial p_t(\mathbf{x})}{\partial t} = -\text{div}\left(\mathbf{f}(\mathbf{x}, t) p_t(\mathbf{x})\right) \Leftrightarrow \frac{d \log p_t(\mathbf{x}(t))}{dt} = -\text{tr}\left(\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}(t)}\right)$$

Solving the continuity equation using the adjoint method is complicated and unstable process.

Flow Matching

$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{x} \sim p_t(\mathbf{x})} \left\| \mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Recap of previous lecture

Let introduce the latent variable $\mathbf{z}$:

$$p_t(\mathbf{x}) = \int p_t(\mathbf{x}|\mathbf{z})p(\mathbf{z})d\mathbf{z}$$

$$\frac{\partial p_t(\mathbf{x}|\mathbf{z})}{\partial t} = -\text{div}\left(\mathbf{f}(\mathbf{x}, \mathbf{z}, t)p_t(\mathbf{x}|\mathbf{z})\right).$$

- $p_t(\mathbf{x}|\mathbf{z})$ is a **conditional probability path**;
- $\mathbf{f}(\mathbf{x}, \mathbf{z}, t)$ is a **conditional vector field**.

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t) \quad \Rightarrow \quad \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{z}, t)$$

Theorem

The following vector field generates the probability path $p_t(\mathbf{x})$.

$$\mathbf{f}(\mathbf{x}, t) = \mathbb{E}_{p_t(\mathbf{z}|\mathbf{x})}\mathbf{f}(\mathbf{x}, \mathbf{z}, t) = \int \mathbf{f}(\mathbf{x}, \mathbf{z}, t)\frac{p_t(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p_t(\mathbf{x})}d\mathbf{z}$$

Tong A., et al. Improving and Generalizing Flow-Based Generative Models with Minibatch Optimal Transport, 2023

Recap of previous lecture

Flow Matching (FM)

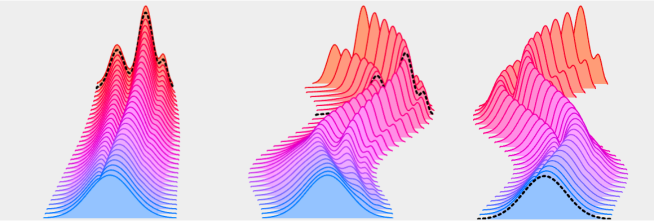
$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{x} \sim p_t(\mathbf{x})} \left\| \mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

Conditional Flow Matching (CFM)

$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})} \mathbb{E}_{\mathbf{x} \sim p_t(\mathbf{x}|\mathbf{z})} \left\| \mathbf{f}(\mathbf{x}, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

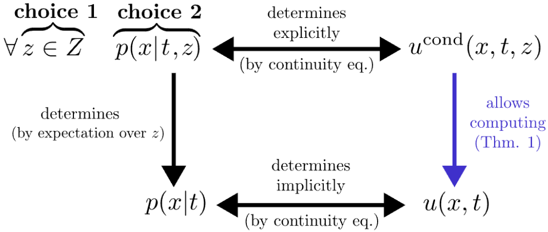
Theorem

If $\text{supp}(p_t(\mathbf{x})) = \mathbb{R}^m$, then the optimal value of FM objective is equal to the optimal value of CFM objective.



Tong A., et al. Improving and Generalizing Flow-Based Generative Models with Minibatch Optimal Transport, 2023

Recap of previous lecture



Constraints

$$p(\mathbf{x}) = \mathcal{N}(0, \mathbf{I}) = \mathbb{E}_{p(\mathbf{z})} p_0(\mathbf{x}|\mathbf{z}); \quad \pi(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_1(\mathbf{x}|\mathbf{z}).$$

- How to choose the conditioning latent variable $\mathbf{z}$?
- How to define $p_t(\mathbf{x}|\mathbf{z})$ which follows the constraints?

Gaussian conditional probability path

$$p_t(\mathbf{x}|\mathbf{z}) = \mathcal{N}\left(\boldsymbol{\mu}_t(\mathbf{z}), \boldsymbol{\sigma}_t^2(\mathbf{z})\right)$$

$$\mathbf{x}_t = \boldsymbol{\mu}_t(\mathbf{z}) + \boldsymbol{\sigma}_t(\mathbf{z}) \odot \mathbf{x}_0, \quad \mathbf{x}_0 \sim p_0(\mathbf{x}) = \mathcal{N}(0, \mathbf{I})$$

image credit: A Visual Dive into Conditional Flow Matching

Recap of previous lecture

Gaussian conditional probability path

$$p_t(\mathbf{x}|\mathbf{z}) = \mathcal{N}\left(\boldsymbol{\mu}_t(\mathbf{z}), \boldsymbol{\sigma}_t^2(\mathbf{z})\right); \quad \mathbf{x}_t = \boldsymbol{\mu}_t(\mathbf{z}) + \boldsymbol{\sigma}_t(\mathbf{z}) \odot \mathbf{x}_0$$

$$\mathbf{f}(\mathbf{x}, \mathbf{z}, t) = \boldsymbol{\mu}_t'(\mathbf{z}) + \frac{\boldsymbol{\sigma}_t'(\mathbf{z})}{\boldsymbol{\sigma}_t(\mathbf{z})} \odot (\mathbf{x} - \boldsymbol{\mu}_t(\mathbf{z}))$$

Conditioning latent variable

Let choose $\mathbf{z} = \mathbf{x}_1$. Then $p(\mathbf{z}) = p_1(\mathbf{x}_1)$.

$$p_t(\mathbf{x}) = \int p_t(\mathbf{x}|\mathbf{x}_1) p_1(\mathbf{x}_1) d\mathbf{x}_1$$

We need to ensure boundary constraints:

$$\begin{cases} p(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_0(\mathbf{x}|\mathbf{z}); (= \mathcal{N}(0, \mathbf{I})) \\ \pi(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_1(\mathbf{x}|\mathbf{z}). \end{cases} \Rightarrow \begin{cases} p_0(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}(0, \mathbf{I}); \\ p_1(\mathbf{x}|\mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1). \end{cases}$$

Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Recap of previous lecture

$$p_0(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}(0, \mathbf{I}); \quad p_1(\mathbf{x}|\mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1).$$

Gaussian conditional probability path

$$p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(\boldsymbol{\mu}_t(\mathbf{x}_1), \boldsymbol{\sigma}_t^2(\mathbf{x}_1)\right); \quad \mathbf{x}_t = \boldsymbol{\mu}_t(\mathbf{x}_1) + \boldsymbol{\sigma}_t^2(\mathbf{x}_1) \odot \mathbf{x}_0.$$

Let consider straight conditional paths

$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_1) = t\mathbf{x}_1; \\ \boldsymbol{\sigma}_t(\mathbf{x}_1) = (1-t). \end{cases} \Rightarrow \quad \begin{cases} p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1, (1-t)^2\mathbf{I}\right); \\ \mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0. \end{cases}$$

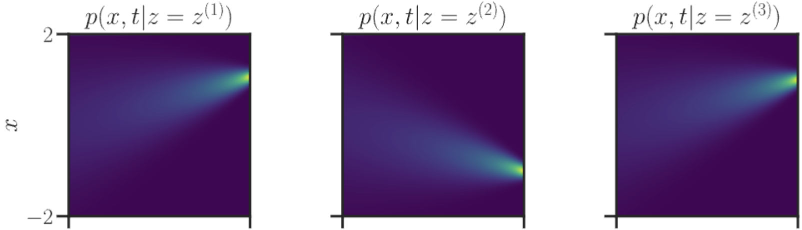


image credit: A Visual Dive into Conditional Flow Matching

Outline

Outline

Conical gaussian paths

$$p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1, (1-t)^2\mathbf{I}\right); \quad \mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0.$$

Conditional vector field

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{x}_1, t) = \boldsymbol{\mu}_t'(\mathbf{x}_1) + \frac{\boldsymbol{\sigma}_t'(\mathbf{x}_1)}{\boldsymbol{\sigma}_t(\mathbf{x}_1)} \odot (\mathbf{x} - \boldsymbol{\mu}_t(\mathbf{x}_1))$$

$$\mathbf{f}(\mathbf{x}, \mathbf{x}_1, t) = \mathbf{x}_1 - \frac{1}{1-t} \cdot (\mathbf{x} - t\mathbf{x}_1) = \frac{\mathbf{x}_1 - \mathbf{x}}{1-t} =$$

$$= \frac{\mathbf{x}_1 - t\mathbf{x}_1 + (1-t)\mathbf{x}_0}{1-t} = \mathbf{x}_1 - \mathbf{x}_0$$

Conditional Flow Matching

$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{z})} \left\|\mathbf{f}(\mathbf{x}, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$

$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{x}_1\sim \pi(\mathbf{x})}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{x}_1)} \left\|\left(\frac{\mathbf{x}_1 - \mathbf{x}}{1-t}\right) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$

$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{x}_1\sim \pi(\mathbf{x})}\mathbb{E}_{\mathbf{x}_0\sim \mathcal{N}(0,\mathbf{I})} \left\|(\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}\left(t\mathbf{x}_1 + (1-t)\mathbf{x}_0, t\right)\right\|^2$$

Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Conditional Flow Matching

$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{x}_1 \sim \pi(\mathbf{x})} \mathbb{E}_{\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})} \left\| (\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

We fit straight lines between noise distribution $p(\mathbf{x})$ and the data distribution $\pi(\mathbf{x})$.

Training

1. Get the sample $\mathbf{x}_1 \sim \pi(\mathbf{x})$.
2. Sample timestamp $t \sim U\{1, T\}$ and $\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})$.
3. Get noisy image $\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \cdot \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \boldsymbol{\epsilon}$.
4. Compute loss $\mathcal{L} = \left\| (\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2$.

Sampling

1. Sample $\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})$.
2. Solve the ODE to get $\mathbf{x}_1$:

$$\mathbf{x}_1 = \text{ODESolve}_f(\mathbf{x}_0, \boldsymbol{\theta}, t_0 = 0, t_1 = 1).$$

Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Flow Matching

$$\mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0$$

- The conditional probability path $p_t(\mathbf{x}|\mathbf{z})$ is an **optimal transport path** from $p_0(\mathbf{x}|\mathbf{z})$ to $p_1(\mathbf{x}|\mathbf{z})$ (in terms of the conditional trajectories straightness).
- The marginal path $p_t(\mathbf{x})$ is not in general an optimal transport path from the standard normal $p_0(\mathbf{x})$ to the data distribution $p_1(\mathbf{x})$.

image credit: https://mlg.eng.cam.ac.uk/blog/2024/01/20/flow-matching.html

Outline

1. Conditional flow matching
   - Conical gaussian paths (continued)
   - Link with score-based models
   - Linear interpolation
   - Rectified flows
2. Latent models
   - Score-based models
   - Autoregressive models
3. The worst course overview

Flow matching vs score-based SDE models

Flow matching probability path

$$p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1, (1-t)^2\mathbf{I}\right); \quad \mathbf{f}(\mathbf{x}, \mathbf{x}_1, t) = \frac{\mathbf{x}_1 - \mathbf{x}}{1-t}$$

Variance Exploding SDE probability path

$$d\mathbf{x} = \sqrt{\frac{d[\sigma^2(t)]}{dt}} \cdot d\mathbf{w}; \quad \Rightarrow \quad \begin{cases} p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(\mathbf{x}_1, \sigma^2_{1-t}\mathbf{I}\right) \\ \mathbf{f}(\mathbf{x}, \mathbf{x}_1, t) = -\frac{\sigma'_{1-t}}{\sigma_{1-t}} \cdot (\mathbf{x} - \mathbf{x}_1) \end{cases}$$

Variance Preserving SDE probability path

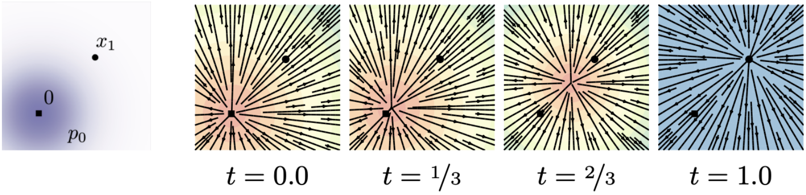
$$d\mathbf{x} = -\frac{1}{2}\beta(t)\mathbf{x}(t)dt + \sqrt{\beta(t)}d\mathbf{w}; \Rightarrow \begin{cases} p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(\alpha_{1-t}\mathbf{x}_1, (1-\alpha^2_{1-t})\mathbf{I}\right) \\ \mathbf{f}(\mathbf{x}, \mathbf{x}_1, t) = \frac{\alpha'_{1-t}}{1-\alpha^2_{1-t}} \cdot (\alpha_{1-t}\mathbf{x} - \mathbf{x}_1) \end{cases}$$

Here $\alpha_t = \exp\left(-\frac{1}{2}\int_0^t \beta(s)ds\right)$.

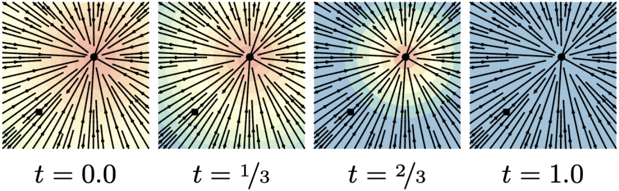
Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Flow matching vs score-based SDE models

Diffusion vector field

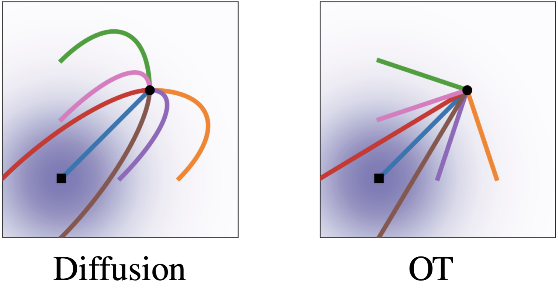


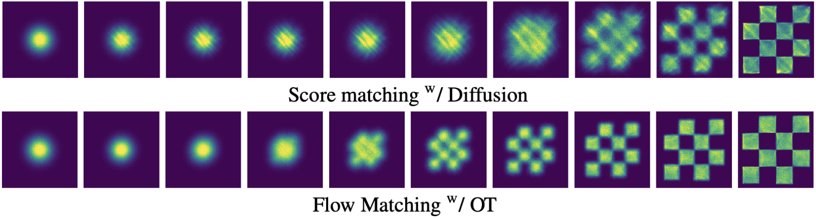
Flow matching vector field



Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Flow matching vs score-based SDE models





Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Outline

Flow Matching

Conditional Flow Matching

$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})} \mathbb{E}_{\mathbf{x} \sim p_t(\mathbf{x}|\mathbf{z})} \left\| \mathbf{f}(\mathbf{x}, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

Let choose $\mathbf{z} = (\mathbf{x}_0, \mathbf{x}_1)$. Then $p(\mathbf{z}) = p_0(\mathbf{x}_0) p_1(\mathbf{x}_1)$.

$$p_t(\mathbf{x}) = \int p_t(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) p_0(\mathbf{x}_0) p_1(\mathbf{x}_1) d\mathbf{x}_0 d\mathbf{x}_1$$

We need to ensure boundary conditions:

$$\begin{cases} p_0(\mathbf{x}) = p(\mathbf{x}) = \mathcal{N}(0, \mathbf{I}); \\ p_1(\mathbf{x}) = \pi(\mathbf{x}). \end{cases} \quad \Rightarrow \quad \begin{cases} p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \\ p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1). \end{cases}$$

Gaussian conditional probability path

$$p_t(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \mathcal{N}\left(\boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1), \boldsymbol{\sigma}_t^2(\mathbf{x}_0, \mathbf{x}_1)\right); \; \mathbf{x}_t = \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1) + \boldsymbol{\sigma}_t^2(\mathbf{x}_0, \mathbf{x}_1) \odot \mathbf{x}$$

Flow Matching

$$\begin{cases} p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \\ p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1). \end{cases} \quad \Rightarrow \quad \begin{cases} \boldsymbol{\mu}_0(\mathbf{x}_0, \mathbf{x}_1) = \mathbf{x}_0, \quad \boldsymbol{\sigma}_0(\mathbf{x}_0, \mathbf{x}_1) = 0 \\ \boldsymbol{\mu}_1(\mathbf{x}_0, \mathbf{x}_1) = \mathbf{x}_1, \quad \boldsymbol{\sigma}_1(\mathbf{x}_0, \mathbf{x}_1) = 0 \end{cases}$$

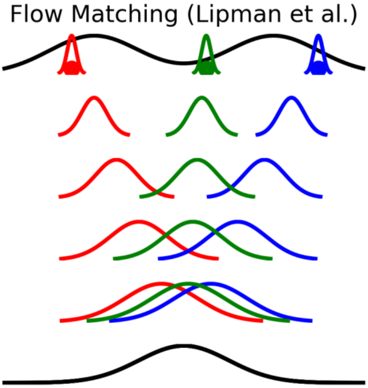
Let consider straight conditional paths

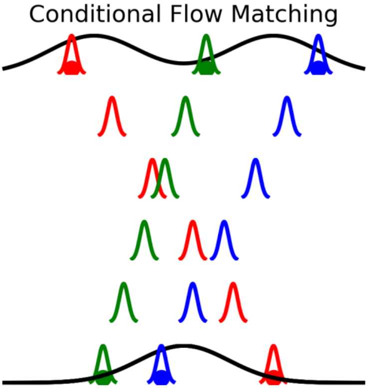
$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1) = t\mathbf{x}_1 + (1-t)\mathbf{x}_0; \\ \boldsymbol{\sigma}_t(\mathbf{x}_0, \mathbf{x}_1) = 0. \end{cases} \quad \Rightarrow \quad \begin{cases} p_t(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1 + (1-t)\mathbf{x}_0 \right. \\ \mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0. \end{cases}$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}_1 - \mathbf{x}_0.$$

Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Flow Matching

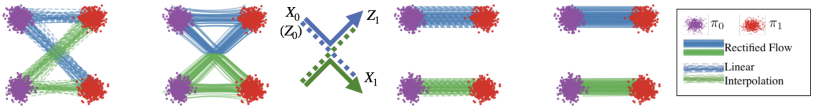




Tong A., et al. *Improving and Generalizing Flow-Based Generative Models with Minibatch Optimal Transport*, 2023

Outline

Rectified flows

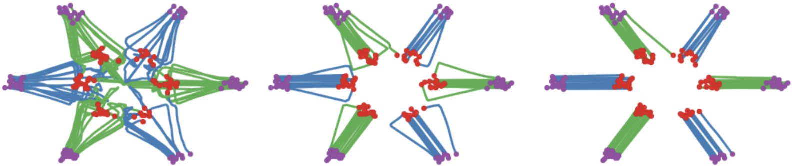


(a) Linear interpolation $X_t = tX_1 + (1-t)X_0$

(b) Rectified flow Z_t induced by (X_0, X_1)

(c) Linear interpolation $Z_t = tZ_1 + (1-t)Z_0$

(d) Rectified flow Z'_t induced by (Z_0, Z_1)



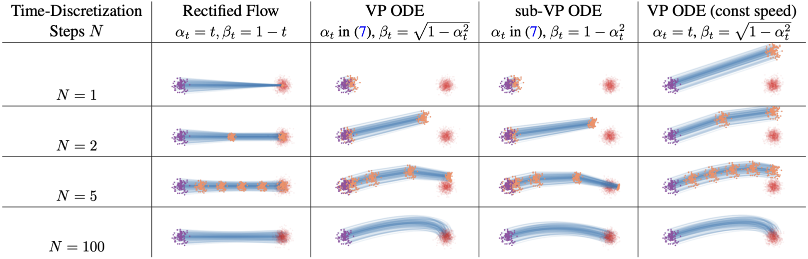
(a)The 1st rectified flow $\boldsymbol{Z}^1$ $\boldsymbol{Z}^1 = \texttt{RectFlow}((X_0, X_1))$

(b) Reflow $\boldsymbol{Z}^2$ $\boldsymbol{Z}^2 = \texttt{RectFlow}((Z_0^1, Z_1^1))$

(c) Reflow $\boldsymbol{Z}^3$ $\boldsymbol{Z}^3 = \texttt{RectFlow}((Z_0^2, Z_1^2))$

Liu X., et al. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow, 2022

Rectified flows



Liu X., et al. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow, 2022

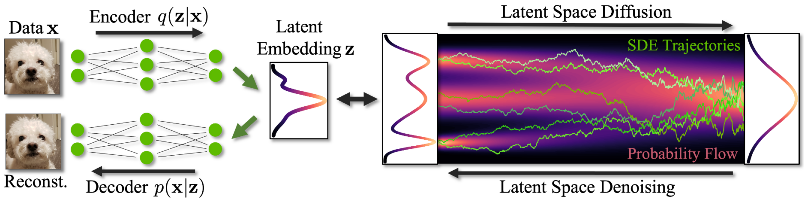
Rectified flows

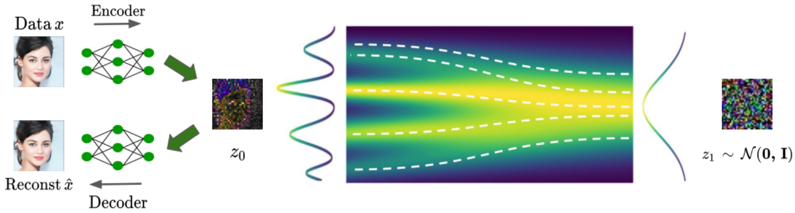
Esser P., et al. Scaling Rectified Flow Transformers for High-Resolution Image Synthesis, 2024

Outline

Outline

Latent models





Dao Q. et al. *Flow Matching in Latent Space*, 2023
NeurIPS 2023 Tutorial: Latent Diffusion Models: Is the Generative AI Revolution Happening in Latent Space?

Outline

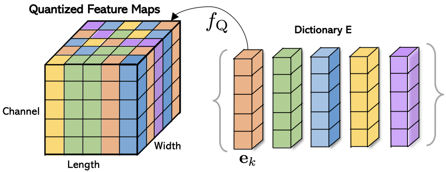
Vector Quantized VAE (VQ-VAE)

Vector quantization

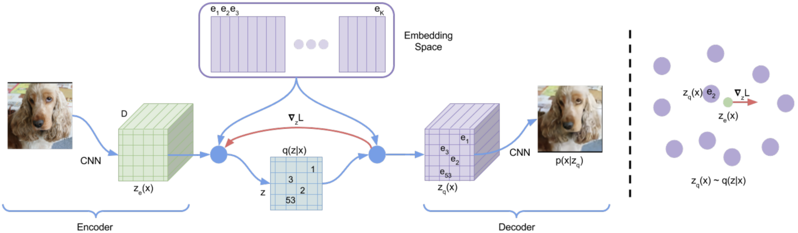
Define the dictionary space $\{\mathbf{e}_k\}_{k=1}^K$, where $\mathbf{e}_k \in \mathbb{R}^C$, K is the size of the dictionary.

$$\mathbf{z}_q = \mathbf{q}(\mathbf{z}) = \mathbf{e}_{k^*}$$

Here $k^* = \arg\min_k \|\mathbf{z} - \mathbf{e}_k\|$.

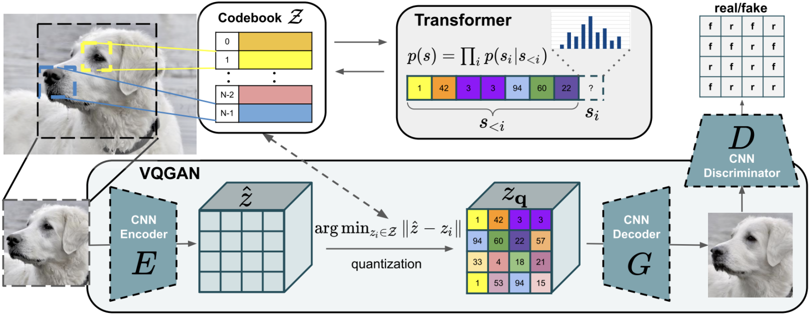


$$\mathcal{L}_{\phi,\boldsymbol{\theta}}(\mathbf{x}) = \log p(\mathbf{x}|\mathbf{z}_q, \boldsymbol{\theta}) - \log K$$



Zhao Y. et al. Feature Quantization Improves GAN Training, 2020
Oord A., Vinyals O., Kavukcuoglu K. Neural Discrete Representation Learning, 2017

Vector Quantized GAN

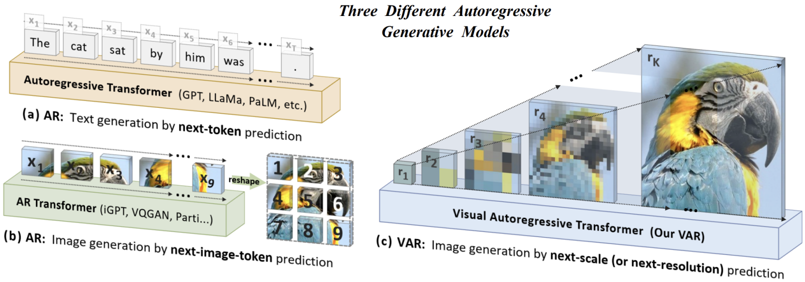


Esser P. et al. Taming Transformers for High-Resolution Image Synthesis, 2020

LlamaGen

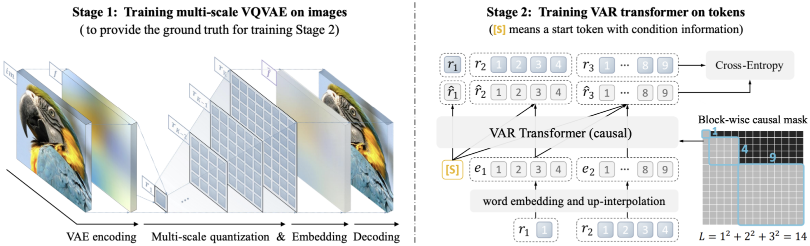
Sun P. et al. *Autoregressive Model Beats Diffusion: Llama for Scalable Image Generation, 2024*

Visual Autoregressive Modeling (VAR)



Tean K. et al. Visual Autoregressive Modeling: Scalable Image Generation via Next-Scale Prediction, 2024

Visual Autoregressive Modeling (VAR)



Tean K. et al. *Visual Autoregressive Modeling: Scalable Image Generation via Next-Scale Prediction, 2024*

Outline

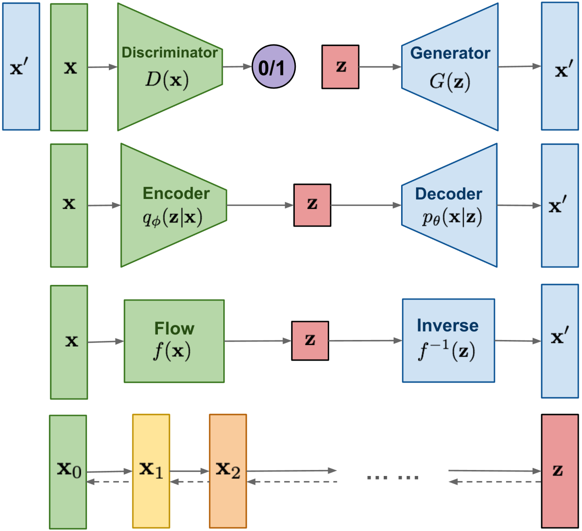
The worst course overview :)

GAN: Adversarial training

VAE: maximize variational lower bound

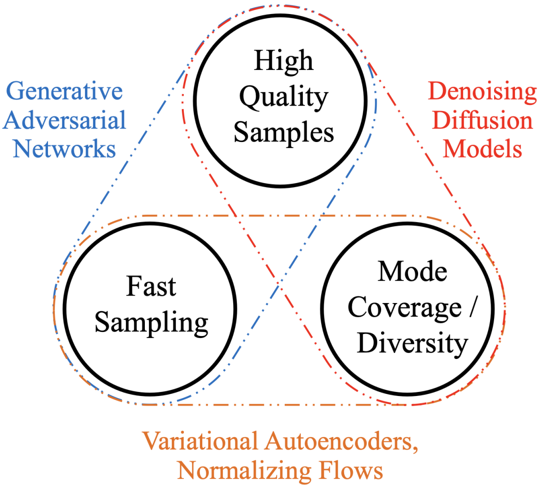
Flow-based models: Invertible transform of distributions

Diffusion models: Gradually add Gaussian noise and then reverse



Weng L. What are Diffusion Models?, blog post, 2021

The worst course overview :)



Xiao Z., Kreis K., Vahdat A. Tackling the generative learning trilemma with denoising diffusion GANs, 2021

The worst course overview :)

Model	Efficient	Sample quality	Coverage	Well-behaved latent space	Disentangled latent space	Efficient likelihood
GANs	✓	✓	✗	✓	?	n/a
VAEs	✓	✗	?	✓	?	✗
Flows	✓	✗	?	✓	?	✓
Diffusion	✗	✓	?	✗	✗	✗

Simon J.D. Prince. Understanding Deep Learning, 2023

Summary

- Conical gaussian paths is the example of the effective FM technique.
-