Deep Generative Models

Lecture 12

Roman Isachenko

Moscow Institute of Physics and Technology
Yandex School of Data Analysis

2025, Autumn

Recap of Previous Lecture

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, t)dt + g(t)d\mathbf{w}$$

Variance Exploding SDE (NCSN)

$$d\mathbf{x} = \sqrt{\frac{d[\sigma^2(t)]}{dt}} \cdot d\mathbf{w}, \quad \mathbf{f}(\mathbf{x}, t) = 0, \quad g(t) = \sqrt{\frac{d[\sigma^2(t)]}{dt}}$$

The variance grows since $\sigma(t)$ is a monotonically increasing function.

Variance Preserving SDE (DDPM)

$$d\mathbf{x} = -\frac{1}{2}\beta(t)\mathbf{x}(t)dt + \sqrt{\beta(t)} \cdot d\mathbf{w}$$

$$\mathbf{f}(\mathbf{x}, t) = -\frac{1}{2}\beta(t)\mathbf{x}(t), \quad g(t) = \sqrt{\beta(t)}$$

The variance is preserved if $\mathbf{x}(0)$ has unit variance.

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

Recap of Previous Lecture

Discrete-Time Objective

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

Continuous-Time Objective

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

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

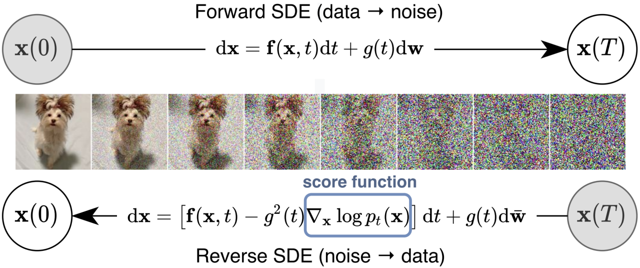
$$q(\mathbf{x}(t)|\mathbf{x}(0)) = \mathcal{N}\left(\mathbf{x}(0)e^{-\frac{1}{2}\int_0^t\beta(s)ds}, \left(1 - e^{-\int_0^t\beta(s)ds}\right)\cdot\mathbf{I}\right)$$

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

Recap of Previous Lecture

Sampling

To sample, solve the reverse SDE using numerical solvers (`SDESolve`).



- Discretizing the reverse SDE gives us ancestral sampling.
- If we use the probability flow ODE instead, then the reverse ODE yields DDIM sampling.

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

Recap of Previous Lecture

Consider ODE dynamics $\mathbf{x}(t)$ in the interval $t \in [0, 1]$ with $\mathbf{x}_0 \sim p_0(\mathbf{x}) = p(\mathbf{x})$, $\mathbf{x}_1 \sim p_1(\mathbf{x}) = p_{\text{data}}(\mathbf{x})$.

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

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.

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

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)

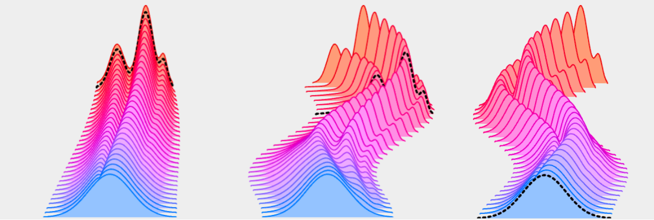
$$\mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{x} \sim p_t(\mathbf{x})} \left\| \mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t) \right\|^2 \to \min_{\boldsymbol{\theta}}$$

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 $\mathrm{supp}(p_t(\mathbf{x})) = \mathbb{R}^m$, then the optimal value of the FM objective equals the optimum for CFM.



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

Outline

1. Conditional Flow Matching

2. Conical Gaussian Paths

3. Linear Interpolation

Outline

Conditional Flow Matching

Theorem

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

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

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

Conditional Flow Matching

Theorem

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

Proof

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

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

Conditional Flow Matching

Theorem

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

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

Proof

$$\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x})} \left\|\mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$

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

$$\mathbb{E}_{p_t(\mathbf{x})}\left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, t)\right] = \int p_t(\mathbf{x})\left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\left(\int p_t(\mathbf{z}|\mathbf{x})\mathbf{f}(\mathbf{x}, \mathbf{z}, t)d\mathbf{z}\right)\right]d\mathbf{x}$$

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

Conditional Flow Matching

Theorem

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

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

Proof

$$\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x})} \left\|\mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$

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

$$\mathbb{E}_{p_t(\mathbf{x})} \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, t)\right] = \int p_t(\mathbf{x}) \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t) \left(\int p_t(\mathbf{z}|\mathbf{x})\mathbf{f}(\mathbf{x}, \mathbf{z}, t)d\mathbf{z}\right)\right] d\mathbf{x} =$$

$$= \int\int p_t(\mathbf{x})p_t(\mathbf{z}|\mathbf{x}) \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, \mathbf{z}, t)\right] d\mathbf{z}d\mathbf{x}$$

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

Conditional Flow Matching

Theorem

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

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

Proof

$$\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x})} \left\|\mathbf{f}(\mathbf{x}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$

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

$$\mathbb{E}_{p_t(\mathbf{x})} \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, t)\right] = \int p_t(\mathbf{x}) \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t) \left(\int p_t(\mathbf{z}|\mathbf{x})\mathbf{f}(\mathbf{x}, \mathbf{z}, t) d\mathbf{z}\right)\right] d\mathbf{x} =$$

$$= \int\int p_t(\mathbf{x})p_t(\mathbf{z}|\mathbf{x}) \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, \mathbf{z}, t)\right] d\mathbf{z}d\mathbf{x} =$$

$$= \mathbb{E}_{p(\mathbf{z})}\mathbb{E}_{p_t(\mathbf{x}|\mathbf{z})} \left[\mathbf{f}_{\boldsymbol{\theta}}^\top(\mathbf{x}, t)\mathbf{f}(\mathbf{x}, \mathbf{z}, t)\right]$$

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

Conditional Flow Matching

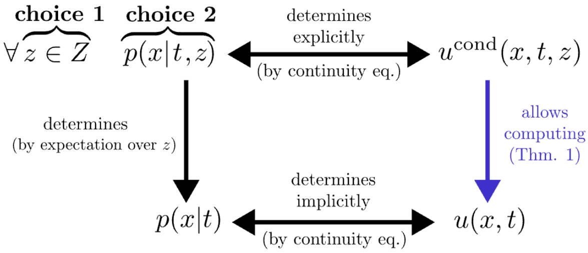
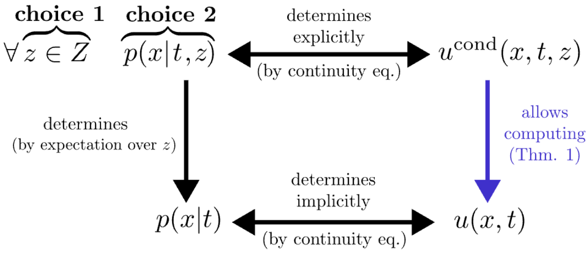


image credit: *A Visual Dive into Conditional Flow Matching*

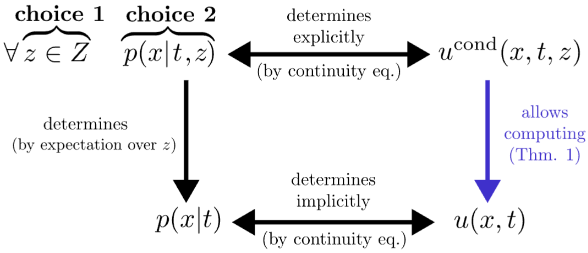
Conditional Flow Matching



- We don't want to directly model $p_t(\mathbf{x})$, since it's complex.
- We've shown it's possible to solve the CFM task instead of the FM task.

image credit: A Visual Dive into Conditional Flow Matching

Conditional Flow Matching



- We don't want to directly model $p_t(\mathbf{x})$, since it's complex.
- We've shown it's possible to solve the CFM task instead of the FM task.
- Let's choose a convenient conditioning latent variable $\mathbf{z}$.
- We'll parametrize $p_t(\mathbf{x}|\mathbf{z})$ instead of $p_t(\mathbf{x})$. It should satisfy the following constraints:

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

image credit: A Visual Dive into Conditional Flow Matching

Conditional Flow Matching

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

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

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

Conditional Flow Matching

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

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

Training

1. Sample a time $t \sim U[0, 1]$ and $\mathbf{z} \sim p(\mathbf{z})$.
2. Draw $\mathbf{x}_t \sim p_t(\mathbf{x}|\mathbf{z})$.
3. Compute the loss $\mathcal{L} = \left\| \mathbf{f}(\mathbf{x}_t, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t, t) \right\|^2$.

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

Conditional Flow Matching

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

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

Training

1. Sample a time $t \sim U[0,1]$ and $\mathbf{z} \sim p(\mathbf{z})$.
2. Draw $\mathbf{x}_t \sim p_t(\mathbf{x}|\mathbf{z})$.
3. Compute the loss $\mathcal{L} = \left\|\mathbf{f}(\mathbf{x}_t,\mathbf{z},t)-\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t,t)\right\|^2$.

Sampling

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

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

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

Conditional Flow Matching

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

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

Conditional Flow Matching

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

Open Questions

Q1 How should we choose the conditioning latent variable $\mathbf{z}$?

Q2 How can we define $p_t(\mathbf{x}|\mathbf{z})$ to enforce the constraints?

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

Conditional Flow Matching

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

Open Questions

Q1 How should we choose the conditioning latent variable $\mathbf{z}$?

Q2 How can we define $p_t(\mathbf{x}|\mathbf{z})$ to enforce the constraints?

Gaussian Conditional Probability Path [Q2]

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

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

Conditional Flow Matching

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

Open Questions

Q1 How should we choose the conditioning latent variable $\mathbf{z}$?

Q2 How can we define $p_t(\mathbf{x}|\mathbf{z})$ to enforce the constraints?

Gaussian Conditional Probability Path [Q2]

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

- There are infinitely many vector fields that generate a particular probability path.
- Let's consider the following dynamics:

$$\mathbf{x}_t = \boldsymbol{\mu}_t(\mathbf{z}) + \boldsymbol{\sigma}_t(\mathbf{z}) \odot \boldsymbol{\epsilon}, \quad \text{with fixed } \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$$

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

Flow Matching

Gaussian Conditional Probability Path

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

Is it possible to derive the expression for $\mathbf{f}(\mathbf{x}_t, \mathbf{z}, t)$?

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

Flow Matching

Gaussian Conditional Probability Path

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

Is it possible to derive the expression for $\mathbf{f}(\mathbf{x}_t, \mathbf{z}, t)$?

Statement

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

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

Flow Matching

Gaussian Conditional Probability Path

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

Is it possible to derive the expression for $\mathbf{f}(\mathbf{x}_t, \mathbf{z}, t)$?

Statement

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

Proof

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{z}, t); \quad \boldsymbol{\epsilon} = \frac{1}{\boldsymbol{\sigma}_t(\mathbf{z})} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{z}))$$

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

Flow Matching

Gaussian Conditional Probability Path

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

Is it possible to derive the expression for $\mathbf{f}(\mathbf{x}_t, \mathbf{z}, t)$?

Statement

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

Proof

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{z}, t); \quad \boldsymbol{\epsilon} = \frac{1}{\boldsymbol{\sigma}_t(\mathbf{z})} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{z}))$$

$$\frac{d\mathbf{x}_t}{dt} = \boldsymbol{\mu}_t'(\mathbf{z}) + \boldsymbol{\sigma}_t'(\mathbf{z}) \odot \boldsymbol{\epsilon} = \boldsymbol{\mu}_t'(\mathbf{z}) + \frac{\boldsymbol{\sigma}_t'(\mathbf{z})}{\boldsymbol{\sigma}_t(\mathbf{z})} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{z}))$$

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

Outline

1. Conditional Flow Matching

2. Conical Gaussian Paths

3. Linear Interpolation

Endpoint Conditioning

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 define our latent variable $\mathbf{z}$.

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

Endpoint Conditioning

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 define our latent variable $\mathbf{z}$.

Conditioning Latent Variable [Q1]

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

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

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

Endpoint Conditioning

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 define our latent variable $\mathbf{z}$.

Conditioning Latent Variable [Q1]

Let us 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 the boundary constraints:

$$\begin{cases} p(\mathbf{x})=\mathbb{E}_{p(\mathbf{z})}p_0(\mathbf{x}|\mathbf{z}); (=\mathcal{N}(0,\mathbf{I})) \\ p_{\text{data}}(\mathbf{x})=\mathbb{E}_{p(\mathbf{z})}p_1(\mathbf{x}|\mathbf{z}) \end{cases}$$

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

Endpoint Conditioning

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 define our latent variable $\mathbf{z}$.

Conditioning Latent Variable [Q1]

Let us 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 the boundary constraints:

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

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

Conical Gaussian Paths

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

Gaussian Conditional Probability Path

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

image credit: A Visual Dive into Conditional Flow Matching

Conical Gaussian Paths

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

Gaussian Conditional Probability Path

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

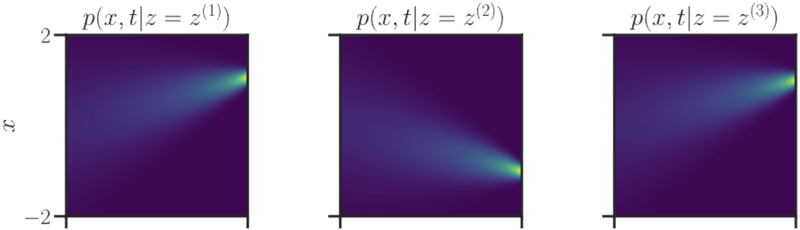


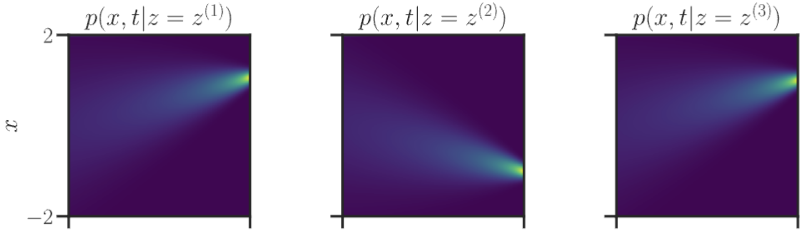
image credit: A Visual Dive into Conditional Flow Matching

Conical Gaussian Paths

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

Gaussian Conditional Probability Path

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



Let's consider straight conditional paths:

$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_1) = t\mathbf{x}_1 \\ \boldsymbol{\sigma}_t(\mathbf{x}_1) = 1 - t \end{cases}$$

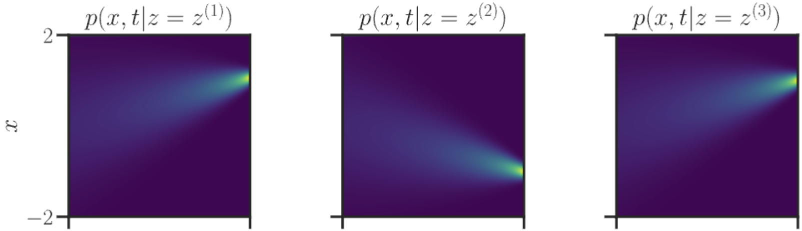
image credit: A Visual Dive into Conditional Flow Matching

Conical Gaussian Paths

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

Gaussian Conditional Probability Path

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



Let's consider straight conditional paths:

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

image credit: A Visual Dive into Conditional Flow Matching

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Conical Gaussian Paths

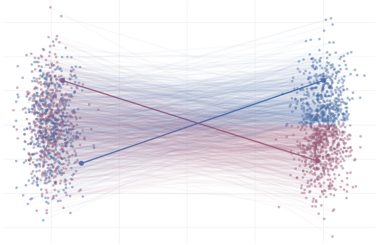
$$p_t(\mathbf{x}|\mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1, (1-t)^2\mathbf{I}\right); \quad \mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0$$

Conditional Vector Field

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

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

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



The conditional vector field $\mathbf{f}(\mathbf{x}_t, \mathbf{x}_1, t)$ defines straight lines between $p_{\text{data}}(\mathbf{x})$ and $\mathcal{N}(0, \mathbf{I})$.

—

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

Conical Gaussian Paths

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

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

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

Conical Gaussian Paths

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

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

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

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

Conical Gaussian Paths

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

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

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

- We fit straight lines between the noise distribution $p(\mathbf{x})$ and the data distribution $p_{\text{data}}(\mathbf{x})$.
- The **marginal** path $p_t(\mathbf{x})$ does not give straight lines.

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

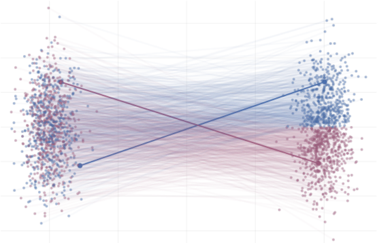
Conical Gaussian Paths

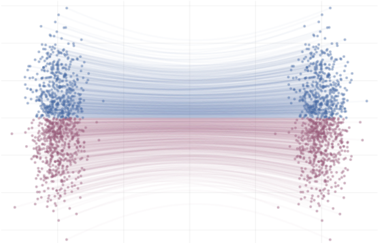
$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{z})}\left\|\mathbf{f}(\mathbf{x},\mathbf{z},t)-\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x},t)\right\|^2=$$

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

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

- We fit straight lines between the noise distribution $p(\mathbf{x})$ and the data distribution $p_{\text{data}}(\mathbf{x})$.
- The **marginal** path $p_t(\mathbf{x})$ does not give straight lines.





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

Conditional Flow Matching

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

Training

1. Sample $\mathbf{x}_1 \sim p_{\text{data}}(\mathbf{x})$.
2. Sample time $t \sim U[0,1]$ and $\mathbf{x}_0 \sim \mathcal{N}(0,\mathbf{I})$.
3. Obtain the noisy image $\mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0$.
4. Compute the loss $\mathcal{L} = \left\|(\mathbf{x}_1-\mathbf{x}_0)-\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t,t)\right\|^2$.

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

Conditional Flow Matching

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

Training

1. Sample $\mathbf{x}_1 \sim p_{\text{data}}(\mathbf{x})$.
2. Sample time $t \sim U[0, 1]$ and $\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})$.
3. Obtain the noisy image $\mathbf{x}_t = t\mathbf{x}_1 + (1 - t)\mathbf{x}_0$.
4. Compute the loss $\mathcal{L} = \left\| (\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t, t) \right\|^2$.

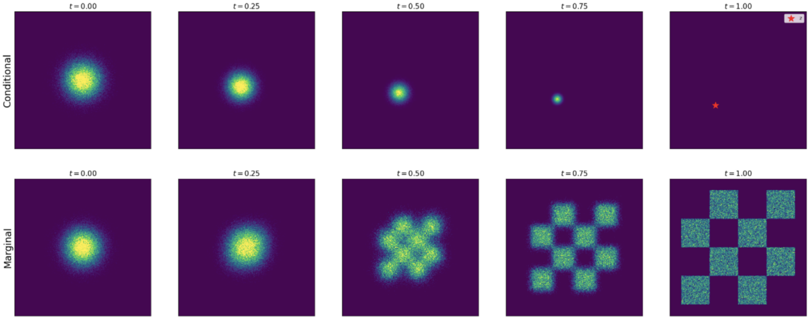
Sampling

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

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

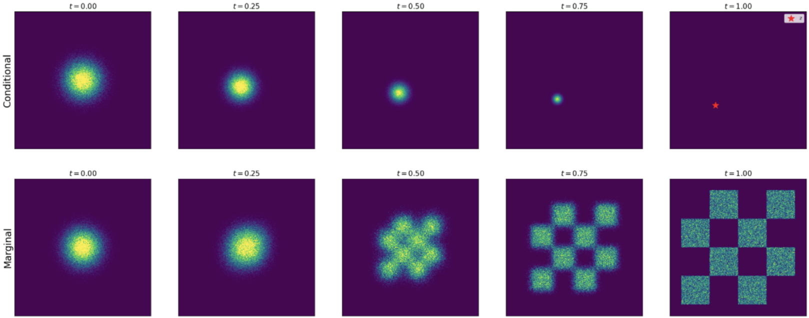
Lipman Y., et al. Flow Matching for Generative Modeling, 2022

Conical Gaussian Paths



Holderrieth P. and Erives E., et al. An Introduction to Flow Matching and Diffusion Models, 2025

Conical Gaussian Paths



- Conical gaussian paths give us the way to construct generative model.
- Now we extend it to image-to-image formulation (mapping between two distinct $p_{\text{data}_1}(\mathbf{x})$ and $p_{\text{data}_2}(\mathbf{x})$).

Holderrieth P. and Erives E., et al. An Introduction to Flow Matching and Diffusion Models, 2025

Outline

1. Conditional Flow Matching

2. Conical Gaussian Paths

3. Linear Interpolation

Pair Conditioning

Conditional Flow Matching

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

Conditioning Latent Variable [Q1]

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

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

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

Pair Conditioning

Conditional Flow Matching

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

Conditioning Latent Variable [Q1]

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

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

We must enforce boundary constraints:

$$\begin{cases} p_0(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_0(\mathbf{x}|\mathbf{z}); \\ p_1(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_1(\mathbf{x}|\mathbf{z}) \end{cases}$$

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

Pair Conditioning

Conditional Flow Matching

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

Conditioning Latent Variable [Q1]

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

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

We must enforce boundary constraints:

$$\begin{cases} p_0(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_0(\mathbf{x}|\mathbf{z}); \\ p_1(\mathbf{x}) = \mathbb{E}_{p(\mathbf{z})} p_1(\mathbf{x}|\mathbf{z}) \end{cases} \Rightarrow \begin{cases} p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0) \\ p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1) \end{cases}$$

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

Linear Interpolation

$$p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \quad p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1)$$

Gaussian Conditional Probability Path

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

image credit: A Visual Dive into Conditional Flow Matching

Linear Interpolation

$$p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \quad p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1)$$

Gaussian Conditional Probability Path

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

Let's consider straight conditional paths:

$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_1) = t\mathbf{x}_1 + (1 - t)\mathbf{x}_0 \\ \boldsymbol{\sigma}_t(\mathbf{x}_1) = \epsilon \end{cases}$$

image credit: A Visual Dive into Conditional Flow Matching

Linear Interpolation

$$p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \quad p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1)$$

Gaussian Conditional Probability Path

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

Let's consider straight conditional paths:

$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_1) = t\mathbf{x}_1 + (1-t)\mathbf{x}_0 \\ \boldsymbol{\sigma}_t(\mathbf{x}_1) = \epsilon \end{cases} \quad \Rightarrow \quad \begin{cases} p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0) \\ p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1) \end{cases}$$

---

image credit: A Visual Dive into Conditional Flow Matching

Linear Interpolation

$$p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0); \quad p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1)$$

Gaussian Conditional Probability Path

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

Let's consider straight conditional paths:

$$\begin{cases} \boldsymbol{\mu}_t(\mathbf{x}_1) = t\mathbf{x}_1 + (1-t)\mathbf{x}_0 \\ \boldsymbol{\sigma}_t(\mathbf{x}_1) = \epsilon \end{cases} \Rightarrow \quad \begin{cases} p_0(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_0) \\ p_1(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \delta(\mathbf{x} - \mathbf{x}_1) \end{cases}$$

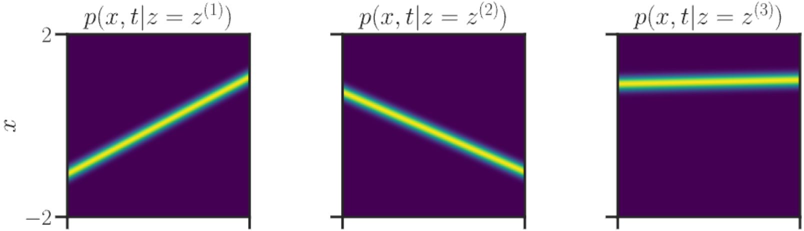


image credit: A Visual Dive into Conditional Flow Matching

Flow Matching: Conical Paths vs. Linear Interpolation

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

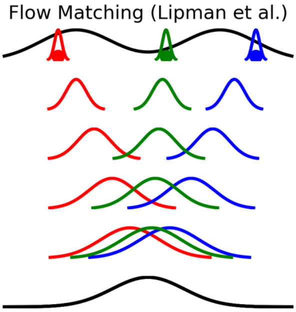
$$\mathbf{z} = (\mathbf{x}_0, \mathbf{x}_1)$$
$$p_t(\mathbf{x}|\mathbf{x}_0, \mathbf{x}_1) = \mathcal{N}\left(t\mathbf{x}_1 + (1-t)\mathbf{x}_0, \epsilon^2\mathbf{I}\right)$$
$$\mathbf{x}_t = t\mathbf{x}_1 + (1-t)\mathbf{x}_0$$

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

Flow Matching: Conical Paths vs. Linear Interpolation

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

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



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

Linear Interpolation

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

Conditional Vector Field

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \boldsymbol{\mu}'_t(\mathbf{x}_0, \mathbf{x}_1) + \frac{\boldsymbol{\sigma}'_t(\mathbf{x}_0, \mathbf{x}_1)}{\boldsymbol{\sigma}_t(\mathbf{x}_0, \mathbf{x}_1)} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1))$$

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

Linear Interpolation

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

Conditional Vector Field

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \boldsymbol{\mu}'_t(\mathbf{x}_0, \mathbf{x}_1) + \frac{\boldsymbol{\sigma}'_t(\mathbf{x}_0, \mathbf{x}_1)}{\boldsymbol{\sigma}_t(\mathbf{x}_0, \mathbf{x}_1)} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1))$$

$$\mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \mathbf{x}_1 - \mathbf{x}_0$$

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

Linear Interpolation

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

Conditional Vector Field

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \boldsymbol{\mu}'_t(\mathbf{x}_0, \mathbf{x}_1) + \frac{\boldsymbol{\sigma}'_t(\mathbf{x}_0, \mathbf{x}_1)}{\boldsymbol{\sigma}_t(\mathbf{x}_0, \mathbf{x}_1)} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1))$$

$$\mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \mathbf{x}_1 - \mathbf{x}_0$$

Conditional Flow Matching

$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{z})}\left\|\mathbf{f}(\mathbf{x}, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$
$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_1)\sim p(\mathbf{x}_0,\mathbf{x}_1)}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{x}_0,\mathbf{x}_1)}\left\|(\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2$$

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

Linear Interpolation

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

Conditional Vector Field

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \boldsymbol{\mu}'_t(\mathbf{x}_0, \mathbf{x}_1) + \frac{\boldsymbol{\sigma}'_t(\mathbf{x}_0, \mathbf{x}_1)}{\boldsymbol{\sigma}_t(\mathbf{x}_0, \mathbf{x}_1)} \odot (\mathbf{x}_t - \boldsymbol{\mu}_t(\mathbf{x}_0, \mathbf{x}_1))$$

$$\mathbf{f}(\mathbf{x}_t, \mathbf{x}_0, \mathbf{x}_1, t) = \mathbf{x}_1 - \mathbf{x}_0$$

Conditional Flow Matching

$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{z})}\left\|\mathbf{f}(\mathbf{x}, \mathbf{z}, t) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2 =$$
$$\mathbb{E}_{t\sim U[0,1]}\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_1)\sim p(\mathbf{x}_0,\mathbf{x}_1)}\mathbb{E}_{\mathbf{x}\sim p_t(\mathbf{x}|\mathbf{x}_0,\mathbf{x}_1)}\left\|(\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}, t)\right\|^2$$

- This yields the same procedure as for conical paths!
- Now, we do not require that $p_0(\mathbf{x})$ is necessarily $\mathcal{N}(0, \mathbf{I})$.

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

Conditional Flow Matching

- This conditioning allows us to transport any distribution $p_0(\mathbf{x})$ to any distribution $p_1(\mathbf{x})$.
- It's possible to apply this approach to paired tasks, e.g., style transfer.

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

Conditional Flow Matching

- This conditioning allows us to transport any distribution $p_0(\mathbf{x})$ to any distribution $p_1(\mathbf{x})$.
- It's possible to apply this approach to paired tasks, e.g., style transfer.

Training Procedure

1. Sample $(\mathbf{x}_0, \mathbf{x}_1) \sim p(\mathbf{x}_0, \mathbf{x}_1)$.
2. Sample time $t \sim U[0, 1]$.
3. Compute the noisy image $\mathbf{x}_t = t\mathbf{x}_1 + (1 - t)\mathbf{x}_0$.
4. Compute the loss $\mathcal{L} = \|(\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)\|^2$.

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

Conditional Flow Matching

- This conditioning allows us to transport any distribution $p_0(\mathbf{x})$ to any distribution $p_1(\mathbf{x})$.
- It's possible to apply this approach to paired tasks, e.g., style transfer.

Training Procedure

1. Sample $(\mathbf{x}_0, \mathbf{x}_1) \sim p(\mathbf{x}_0, \mathbf{x}_1)$.
2. Sample time $t \sim U[0, 1]$.
3. Compute the noisy image $\mathbf{x}_t = t\mathbf{x}_1 + (1 - t)\mathbf{x}_0$.
4. Compute the loss $\mathcal{L} = \|(\mathbf{x}_1 - \mathbf{x}_0) - \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)\|^2$.

Sampling

1. Sample $\mathbf{x}_0 \sim p_0(\mathbf{x})$.
2. Solve the ODE to obtain $\mathbf{x}_1$:

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

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

Stable Diffusion 3: Scalable Flow Matching

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

Summary

- Conditional flow matching makes the FM objective tractable.
- Conical Gaussian paths use endpoint conditioning $\mathbf{z} = \mathbf{x}_1$ and serve as an effective FM technique.
- Linear Interpolation paths use pair conditioning $\mathbf{z} = (\mathbf{x}_0, \mathbf{x}_1)$ and yields the same procedure, but is more general (suitable for paired tasks).