Temporal Difference Flows [FPT+25]

Kellen Kanarios

July 3, 2025

Outline

1. Diffusion Recap
   - Continuity Equation
2. Flow Matching
   - Flow vs. Score
   - Learning the Flow
3. Application to Reinforcement Learning
   - Successor Measures

Continuity Equation

Def.

$$\mathrm{d}X_t = v(t, X_t)\mathrm{d}t$$

Eqn.

$$\frac{\partial}{\partial t}\boldsymbol{p} + \nabla_x(\boldsymbol{p} \cdot \boldsymbol{v}) = 0 \tag{1}$$

Proof sketch.

$$\frac{\partial}{\partial t}(p_t(X_t)) = \frac{\partial}{\partial t}p_t(X_t)\mathrm{d}X_t \tag{2}$$

$$x = y \tag{3}$$

$$x = y \tag{4}$$

$$x = y. \tag{5}$$

Diffusion vs. Flow Matching

Diffusion Models

- Learn the score $\nabla \ln p_t$
- Corresponds to one particular choice of vector field

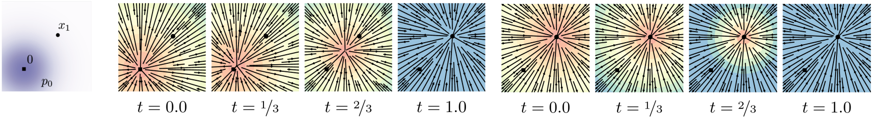
Flow Models

- Directly learn the vector field v_t
- Corresponds to one choice of "flow"

Why Flows?

Anecdotally

- *The OT path's conditional vector field has constant direction in time and is arguably simpler to fit with a parametric model. [LCB+23]*
- *The deterministic nature of ODEs equips flow-matching methods with simpler learning objectives and faster inference speed [ZPLE25]*



Conditional score

Conditional vector field

Naive Loss + Problems

Defs.

- p_t
- $u(t, \cdot)$

Flow Model Loss

$$\mathcal{L}_{\text{FM}}(\theta) = \mathbb{E}_{t, p_t(x)} \| v_\theta(t, x) - u(t, x) \|^2$$

Naive Loss + Problems

Defs.

- p_t
- $u(t, \cdot)$

Flow Model Loss

$$\mathcal{L}_{\text{FM}}(\theta) = \mathbb{E}_{t, p_t(x)} \| v_\theta(t, x) - u(t, x) \|^2$$

Question

How to sample from p_t, or compute $u(t, \cdot)$?

Naive Loss + Problems

Defs.

- p_t
- $u(t, \cdot)$

Flow Model Loss

$$\mathcal{L}_{\text{FM}}(\theta) = \mathbb{E}_{t, p_t(x)} \left\| v_\theta(t, x) - u(t, x) \right\|^2$$

Question

How to sample from p_t, or compute $u(t, \cdot)$?

Answer

We don't have to!!!!

Conditioning Trick

More defs.

- $\frac{\mathrm{d}}{\mathrm{d}t}\psi_t(x) = u(t, \psi_t(x))$

Equivalence of loss functions

Example: Optimal Transport

For $\psi_t(x) = \mu_t(x_1) + x\sigma_t(x_1)$, we consider

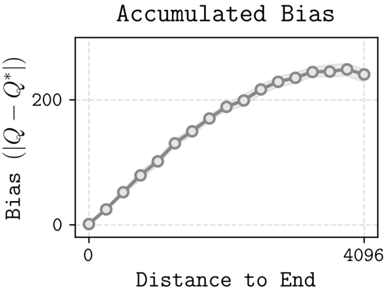
$$\mu_t = tx_1, \quad \sigma_t(x) = 1 - (1-t)\sigma_{\min}.$$

Recall

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t(x) = u(\psi_t(x) \mid X_1).$$

Long Horizon Problems Are Hard

[Par25] ***Seohong Park***; *Q learning is not yet scalable*



$$\mathbb{E}_{(s,a,r,s')\sim\mathcal{D}}\left[\left(Q_\theta(s,a) - \underbrace{\left(r + \gamma \max_{a'} Q_{\bar{\theta}}(s',a')\right)}_{\texttt{Biased}}\right)^2\right].$$

Successor Measure

References

Jesse Farebrother, Matteo Pirotta, Andrea Tirinzoni, Rémi Munos, Alessandro Lazaric, and Ahmed Touati.
Temporal Difference Flows, March 2025.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le.
Flow Matching for Generative Modeling, February 2023.

Seohong Park.
Q-learning is not yet scalable, June 2025.

Chongyi Zheng, Seohong Park, Sergey Levine, and Benjamin Eysenbach.
Intention-Conditioned Flow Occupancy Models, June 2025.